Aalto University
School of Business

# Disaggregation of Finnish Rental Markets – Empirical Evidence from Helsinki Metropolitan Area

Master's Thesis
Joonas Bienek
Fall 2017
Finance

Approved in the Department of Finance __ / __20___ and awarded the grade

___

**Aalto University**
**School of Business**

**Aalto University, P.O. BOX 11000, 00076 AALTO**
**www.aalto.fi**

**Abstract of master's thesis**

**Author** Joonas Bienek

**Title of thesis** Disaggregation of Finnish Rental Markets – Empirical Evidence from Helsinki Metropolitan Area

**Degree** Master of Science (Economics and Business Administration)

**Degree programme** Master’s Programme in Finance

**Thesis advisors** Markku Kaustia

**Year of approval** 2017 **Number of pages** 71 **Language** English

The purpose of this study is to examine whether the rental market in Helsinki metropolitan area is formed of distinct segments of one-, two- and three-room apartments. Evaluation is conducted by comparing the determinants of net rental yield between these distinct apartment subgroups. Results are validated by examining the statistical significant model and coefficient difference between subgroup regression models.

The novel dataset employed in this study is received from multiple sources. Data of rental advertisements was received from Oikotie. Complementary data regarding apartment transaction prices and apartment characteristics was received from KVKL, data regarding apartment supply was received from Statistics Finland and the rest was obtained from public sources. Independent variables employed in this study to examine the net rental yield differences between the apartment subgroups cover apartment characteristics, location related variables, neighborhood related variables and a variable expressing the supply of new apartments in a postal code area.

To examine the disaggregation of rental markets, the aggregated rental function is compared to distinct double-log regressions by apartment type. Moreover, the statistical significance between the regressions models and rental yield determinants is explored with a Chow-test and Tiao-Goldberger F-test.

The results imply that the apartment rental market in Helsinki metropolitan area is disaggregated and the important rental yield determinants differ between apartment types. The distinct regressions for different apartments are statistically significantly different from each other and from the aggregated model. Key variables with different impact on rental yield between the apartment types, one-, two- and three-room apartments, are the size of apartment, the building year of the apartment, the apartment supply in postal code area and the proximity of closest university and metro station.

**Keywords** Real Estate, Apartments, Rental market, Rental yield, Submarket, Disaggregated

Aalto University School of Business

Aalto-yliopisto, PL 11000, 00076 AALTO
www.aalto.fi

**Tutkielman tiivistelmä**

**Tekijä** Joonas Bienek

**Tutkielman nimi** Suomen vuokramarkkinoiden segmentointi asuntotyypeittäin – empiirisiä havaintoja pääkaupunkiseudulta

**Tutkinto** Kauppatieteiden maisteri

**Koulutusohjelma** Rahoitus

**Työn ohjaaja(t)** Markku Kaustia

**Hyväksymisvuosi** 2017 **Sivumäärä** 71 **Kieli** Englanti

Tutkielman tavoitteena on tutkia muodostuvatko yksiöiden, kaksioiden ja kolmioiden vuokramarkkinat pääkaupunkiseudulle erilaisista tekijöistä. Tämä vertailu tehdään vertailemalla näiden asuntotyyppien nettovuokratuottoon vaikuttavia muuttujia. Tulokset validoidaan varmentamalla vuokrafunktioiden tilastollinen merkitsevyys ja tutkimalla mitkä spesifit muuttujat aiheuttavat tämän eroavaisuuden nettovuokratuoton muodostuksessa asuntotyyppien välillä.

Akateemisessa kontekstissa aikaisemmin käyttämätön aineistoni on koottu useista eri lähteistä. Aineisto koskien vuokratasoja ja asuntojen ominaispiirteitä on saatu Oikotieltä, myyntihintoja koskevat tilastot on saatu Kiinteistövälittäjien keskusliitolta, asuntojen tarjontaan liittyvä postinumerokohtainen aineisto on saatu Suomen Tilastokeskukselta ja loput on kerätty julkisista lähteistä. Tutkielmani selittävät muuttujat kuvaavat asuntojen ominaispiirteitä, asuntojen sijaintia, asuntojen ympäristöä ja muita vuokratuottoon vaikuttavia tekijöitä, kuten esimerkiksi asuntojen tarjonta postinumeroalueella.

Ensimmäinen askel hypoteesini testaamiseen on asuntojen nettovuokratuotto funktioiden määrittäminen logaritmisella lineaarisella regressiolla ja muuttujien myötävaikutuksen vertaileminen asuntotyyppien välillä. Tätä vaihetta seuraa funktioiden tilastollisesti merkittävän eroavaisuuden tutkiminen Chow-testilla. Lopuksi tutkin mitkä muuttujista eroavat tilastollisesti merkittävästi mallien välillä Tiao-Goldberger testillä.

Tulokset implikoivat sitä, että Suomessa on eriytyneet vuokramarkkinat yksiöille, kaksioille ja kolmioille. Estimoimani mallit eri asuntotyypeille ovat tilastollisesti merkittävästi erilaiset. Tiao-Goldberger testi validoi nämä tulokset. Suurimmat eroavaisuudet vuokramarkkinoiden välillä tulevat asuntojen koon, asunnon rakennusvuoden ja yliopisto kampuksen, sekä metro-aseman läheisyyden vaikutuksesta asuntojen vuokratuottoon.

**Avainsanat** Kiinteistöt, Asunnot, Vuokramarkkinat, Vuokratuotto, Segmentointi

**Table of Contents**

**List of Tables**



## List of Figures

# 1. INTRODUCTION

Apartment characteristics differ depending on whom they are built for – one-room apartments are often quite compact and provide only the necessary amenities, being designed for single-dwellers. Two-room apartments offer extra living space and amenities and are hence comfortable for multiple dwellers. Finally, three-room apartments provide further living space, amenities and comfort for additional family members. Despite the clear differences between the diverging characteristics of different apartment types and profiles of possible tenants, in both academic literature and practical investment decisions, apartments are often treated as a single asset-class. Assessing the real estate apartment market as a sum of distinct apartment subgroups instead of homogenous mass is still a research field largely unexplored.

One of the first papers highlighting the financial benefits of diversification within real estate asset class was written by Hartzell et al. (1986). After this initial study, majority of the academic literature regarding real estate diversification has focused on either geographic analysis of the market (see e.g. Williams 1996) or specific property-type market research (see e.g. Fehribach et al., 1993). Moreover, as discussed by Clapp et al. (1992) in their study examining distinct apartment submarkets within the office real estate segment, majority of the studies regarding the rental markets are based on aggregate samples instead of examining the possible heterogeneous submarkets within the real estate apartment market.

There are, however, few examples of studies contemplating the existence of distinct apartment submarkets. Allen et al. (1995) discovered that residential rental equations were significantly different for the following dwelling types: apartments, condominiums and detached single-family residences. Similarly, Black et al. (1997) discovered differing rental equations between industrial distribution and manufacturing property types. Wolverton et al. (1999), conducted one of the first studies examining disaggregation in apartment rental market between one-room, one-bath units and two-room, one-bath units. They discovered that the rental equations between the apartment types were statistically significantly different between the submarkets, hence solidifying the hypothesis of disaggregated apartment rental markets.

In Finnish context, the past studies have focused solely on examining the nuances of aggregated market (see e.g. Laakso, 1997; Terho and Moilanen 2010). The examination of apartment submarkets and hence the identification of real estate diversification possibilities between apartment types of areas has been non-existent. Despite the lack of academic research on the subject, Finnish population continues to allocate capital into real estate assets. As of 2016, the Bank of Finland estimated that up to 53 percent of the median household's wealth is tied to owner-occupied dwelling. However, in the Helsinki metropolitan area (HMA), the picture is slightly different. In this area, rental dwelling is relatively more popular than in the rest of Finland. In 2016, on average 32 percent of Finns were living on rented apartments, whereas the number was 49 percent in Helsinki area.[1] Thus, given the national importance of apartment and rental markets, profoundly understanding the segments of rental markets in Finland would be economically highly important. Proper segmentation of the rental market would lead into more sophisticated capital allocation within real estate market as both retail and institutional investors could invest in assets with optimal characteristics, regardless of the apartment type.

Due to the lack of academic research on disaggregated rental markets and the relative importance of real estate as an asset class in Finland, this study aims to extend the empirical research on Finnish rental market. This study contributes to the existing academic literature in three ways. Firstly, this is the first study, to my best knowledge, examining the hypothesized disaggregation of the Finnish rental market instead of treating it as an aggregate market. Secondly, unlike the previous studies on Finnish rental markets (see e.g. Terho and Moilanen, 2010), I had access to actual transaction prices instead of advertised prices to calculate net rental yields for the rental apartments. Operating with the realized transaction prices instead of ask prices should increase the accuracy of the models and reduce biases related to differences between ask and realized apartment prices. Finally, this is the first study incorporating the apartment supply on postal code level as a determinant explaining the formation of net rental yield.

Determinants of net rental yield are assessed with a regression model including property's characteristics, location related variables, variables related to apartments neighbourhood and other variables often linked to rental yield in empirical research. Regression analyses are computed for both aggregated and disaggregated sample of one-, two- and three-room apartments. The lack of universal determinants for net rental yield across all apartment types seems intuitive and would allow rental investors to make even more sophisticated investment decisions. Hence, revealing the magnitude the net rental yield determinants have in different apartment types has large economic implications for both retail and institutional investors.

[1] Statistics Finland, www.stat.fi

Initial regression analysis is followed with a Chow-test and Tiao-Goldberger F-test to assess the statistically significant difference between subgroup regressions.

This study is conducted with a novel dataset received from both Oikotie and KVKL (English: Central Federation of Finnish Real Estate Agencies). Oikotie is Finland´s most popular apartment sales and rental website whereas KVKL is the umbrella organization for all Finnish real estate agencies. The data received from Oikotie contains all rental advertisements in HMA listed to their website between 2012 and 2016. This data contains a large amount of apartment characteristics from size to whether the apartment has a balcony or not. The KVKL dataset, on the other hand, contains all transactions made by Finnish real estate agents between 1998 and 2016. For the purposes of this study, matching dataset of apartment transactions in HMA is hand-collected for the time-period 2012-2016. This dataset contains information regarding the geographical location and the characteristics of the apartments along with the actual transaction price paid for the apartment. In this study, I match these two datasets to create a comprehensive picture of HMA rental markets with a non-biased dataset. Furthermore, I received supply side related variables from Statistics Finland, coordinates of apartments from Google Maps and coordinates of university campuses, metro and railway stations from HSL (Helsinki Region Transport authority). This additional data is merged to the previously matched data of rental apartments with actual transaction prices, based on postal code areas or coordinates.

My results clearly indicate that the apartment rental markets are disaggregated between one-, two- and three-room apartments in HMA. The determinants in the double log-regressions models for the distinct apartment subgroups have varying economical and statistical significances. Moreover, examination of the slopes of one-, two- and three-room apartment regressions with Chow-test imply that net rental yield functions are statistically significantly different between all subgroup pairs. Furthermore, Tiao-Goldberger F-test revealed that the following coefficients are causing the statistical differences between the functions: apartment’s size, the building year, the existence of sauna, is the apartment building on a owned lot, construction supply in postal code area and the proximity of university campus and metro station. Thus, my empirical results imply that simply treating the real estate apartment’s asset class as a homogenous mass might lead to un-optimal investment decisions.

The structure of this study is following. Firstly, I briefly discuss the development and characteristics of Finnish apartment and rental markets to lay out the context for this study. Secondly, I discuss the related academic literature in detail. I focus on especially on literature regarding the rental yield composition, apartment submarkets and the determinants of rental yield. Thirdly, I discuss the methodology employed in my study. Fourthly, I explain the data used in this study and the various sources from which I have obtained it. Fifthly, I discuss my results regarding the disaggregation of Finnish rental markets. Finally, I conclude my study.

# 2. PRACTICAL BACKGROUND

The purpose of this chapter is to create a coherent picture of the Finnish apartment and rental markets. Firstly, I explain how the Finnish apartment and rental markets have developed over the past decades. Secondly, I open the structure and characteristics of the current property market in Finland. Finally, I briefly discuss the consumption preferences of Finnish dwellers in the context of owner-occupancy and rental markets. Due to the illiquid nature of real estate, emphasis of this chapter is on longer time-period to unveil the impact of structural and political changes in the apartment and rental markets. Both HMA specific and overall Finnish trends and preferences are discussed to form a comprehensive context for this study.

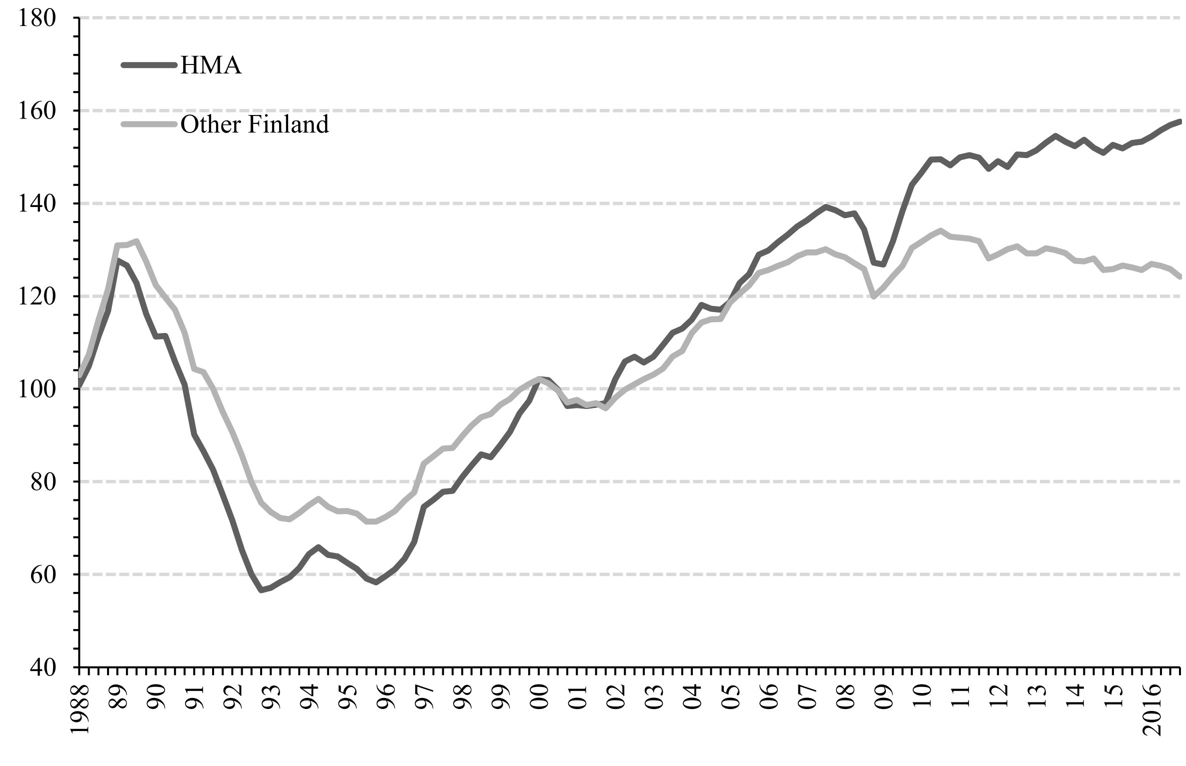
## 2.1 Development of Finland's housing market

Oikarinen (2007) provides an extensive perspective to the development of Finnish housing market; the largest factors causing the past market development are large housing market deregulation reforms combined with the entrance of international investors to the market. Opening property markets in the 1970s for international investors caused housing prices to become more volatile than in the past when capital in the market was scarce. First symptom of this phenomenon was the high inflation and related nominal price increases post-oil crisis in the 1970s. However, the prices of apartments on real terms started appreciating only from the 1980s. The key reason for this was the aforementioned gradual opening of capital markets on the other hand, and good employment situation and stable economic growth on the other hand, as loans became more accessible to retail clients (Huovari et al. 2006). Moreover, the easy access to credit, i.e. abolishment of average lending rates by Bank of Finland, coupled with decreasing down payment ratios led to huge growth in credit and a housing market boom in the 1980s (Oikarinen, 2007). In the 1990s, this growth phase ended abruptly to depression period, which led to further property market reforms.

Figure 1 showcases the development of real housing prices during the past four decades. As previously discussed, the first overheating of housing markets was sparked by the opening of financial markets in 1984. The following price depreciation was largely caused by the recession starting in the 1990s. Furthermore, the figure illustrates how in general HMA price levels have stayed on a lower level compared to the rest of Finland until the burst of the internet bubble in the year 2000. After the internet bubble, HMA apartment prices have been more resilient than the general apartment prices to price depreciation. For example, this development is imminent in the years following the financial crisis in 2008. After 2008, the apartment markets have begun to diverge in terms of price; apartment values in HMA have increased steadily whereas the price index for the rest of Finland seems to be on a downwards spiral.

**Figure 1 - Price index for Finnish apartments, Q1/1988 – Q4/2016, 2000 =100, real values**

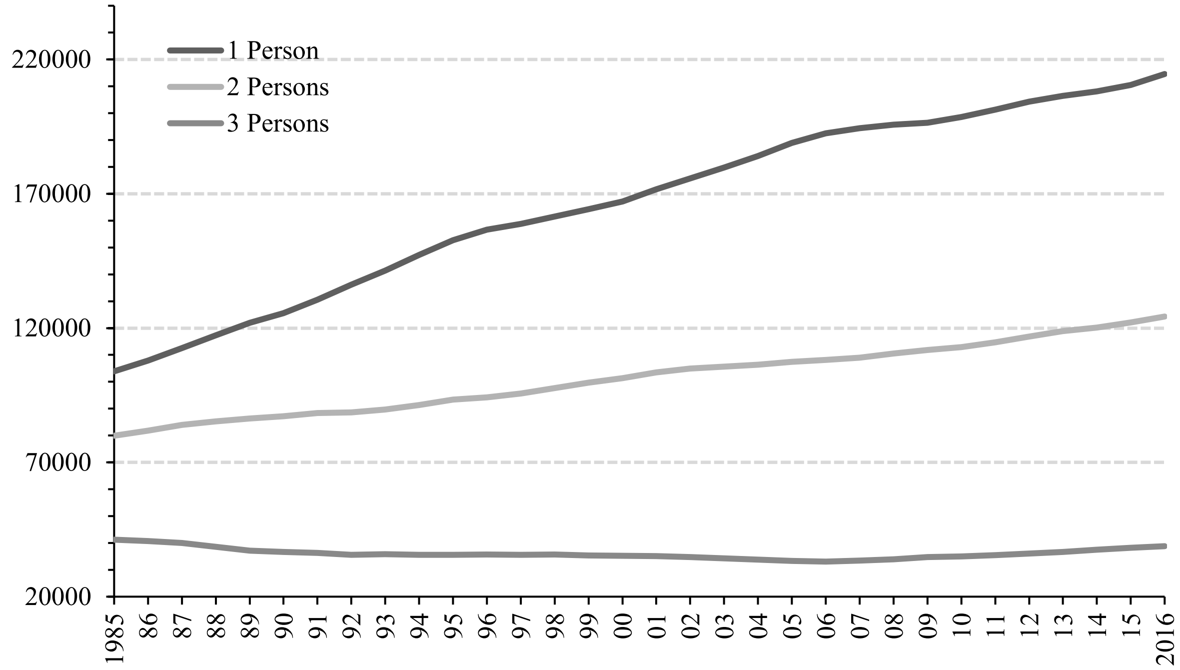
This figure illustrates the price appreciation of apartments in HMA and the rest of the Finland. By definition, rest of the Finland is whole Finland's price appreciation excluding HMA's price appreciation. The timeframe is from Q1/1988 to Q4/2016 and the base year of the index is set at the year 2000. The graph illustrates partly the spike in apartment prices after market deregulation in 1986, visible between the years 1988-1989. This is followed by the rapid price decline in 1990s and finally the currently diverging price levels between HMA and rest of Finland – a phenomenon started after the global financial crisis in 2008.




In the context of disaggregated rental markets, it is also noteworthy to examine how the number of different sized households have developed over the past years. Figure 2 highlights this development of household unit size from 1985 to 2016 in HMA**.** First clear observation from the figure is how one- and two-person households have increased in popularity steadily from 1985 whereas the number of larger households has stayed relatively stable throughout the whole-time period.

**Figure 2 - Development of household unit size in Finland 1985-2016**

This figure shows how the household composition has developed in Finland between the years 1985 and 2016. The darkest line represents one-person households, the lightest line households with two-persons and the medium-dark line households with three-persons. Household composition implies how many persons are living in same apartment, not the type of apartment they live in. Figure indicates that since the tracking started in 1985, more and more Finns choose to live in one - or two-person households whereas the number of three-person households has stayed relatively stable.




The two clear trends causing this development in household size are slowing population growth and behavioral shift towards single dwelling. Firstly, regarding the population growth, the fertility rate and thus the demand for large apartments has decreased steadily in Finland. According to Statistics Finland, the total fertility rate of Finnish women living in Uusimaa region, i.e. the region containing Helsinki, Vantaa, Espoo has decreased from 1.58 in 2013 to 1.51 in 2015. Secondly, the popularity of single dwelling can be explained by behavioral shift, e.g. the need of youth to live independently[2]. Moreover, the demographic trend of decreasing population under 15 years old and the simultaneous increase in 65-year old's increases the popularity of single dwelling.[2]

As a product of these factors, Finns are living in smaller and smaller apartments as the household size keeps decreasing due to reduced fertility and preferences for solo dwelling. Furthermore, these macro-trends affect the demand and thus the price of rental apartments in

[2] www.ptt.fi

the submarkets of one-room and to smaller extent two-room apartments – likely simultaneously driving the rental levels in these apartments higher levels.

**2.2 Structure of Finland’s apartment and rental markets**

Finnish apartment and rental markets can be classified into two main segments: the privately financed sector and the subsidized sector of publicly owned or regulated apartments. The private sector has been able to freely set the prices and rents after rent control system was discontinued in 1995. Only exception to the rule is that rents in existing contracts cannot be increased without including indexation clauses in the rental contracts. As a product of the principal of free transacting between the rental agreement parties, KTI (2017) considers Finnish rental market among the most liberal in the world.

In the subsidized sector, the government entities ultimately act as the price and rent setters of apartment. Hence, the prices of subsidized apartments are often set below market prices to cater for the whole population. Even as the target customers between private and public housing segments vary, it is important not to completely neglect the impact how subsidized sectors dynamics impact the private housing sector. DiPasquale and Wheaton (1996), discovered that the construction of subsidized housing might have adverse impact for the demand of privately owned rental units. Moreover, this increased supply of public housing leads to lower rental rates and apartment prices and hence decreases the financial incentives for private construction projects in the area. In the context of this study, due to unavailable data regarding the rental contracts or apartments prices for subsidized sector in my data sources, I need to omit this sector from the scope my study.

Regarding the size of the private and public housing markets, Niemi et al. (2011), estimated the value of Finnish housing stock to be 230 billion euros in their study regarding Finland’s populations wealth. Moreover, according to Oikarinen (2007), approximately half of the rental homes are owned by institutional investors or by government owned entities. Largest institutional investors operating in Finland are pension funds, insurance accompanies and not-for-profit foundations. However, during past few years Finnish companies previously acting solely in the subsidized housing sector, such as VVO and Sato, have increased their presence in the private sector. Furthermore, the Finnish housing market has seen the formation of property funds focusing on rental apartments in the 2000s and the initial public listing of first Finnish rental property investment company, Orava.

When it comes to market characteristics, Statistics Finland's statistics imply that in 2015 owner-occupancy was clearly the most popular form of living in Finland, with 64 percent of dwellers living in owner-occupied apartments. Furthermore, only 32 percent of Finnish dwellers are choosing rental dwelling. According to Carliner and Marya (2016), this level is clearly lower than in comparable European countries. For example, they discovered that Germany's rental level stands at 55 percent, UK's at 36 percent and Sweden's at 38 percent. Rest of the apartment market in Finland is comprised of "right-of-occupancy" dwellings. The high share of owner-occupancy dwellings is likely motivated by tax-incentives targeted to home ownership in Finland**.** For example, according to Finnish tax authorities as of 2017, Finns are able to tax-deduct 45 percent of their yearly mortgage interest payments from their capital income.[3] There are no tax-incentives for rental dwelling.

### 2.3 Dwelling preferences of Finnish consumers

In this section, I focus on reviewing literature regarding consumers dwelling preferences in Finland. The emphasis in review is on studies focusing on HMA, however also relevant studies encompassing overall Finnish market are discussed. I start by reviewing the key macro-trend influencing Finland's dwellers decisions making process; urbanization. Secondly, I review Juntto's (2007) key findings about the general preferences of Finnish dwellers. Finally, I discuss the main characteristics and preferences of consumers in HMA.

Urbanization in Finland has its' roots in the ever-globalizing trade and increasing education of Finnish population. Finnish towns were originally built close to local factories, e.g. paper mills in the rural areas of Finland. However, increasing competition from global markets has decreased the demand for the produced products and led to organization restructurings, often including layoffs and manufacturing plant relocations. Due to limited employment options beyond the local factories in rural areas, laid off personnel were in practice forced to relocate to seek further employment. Often this relocation led to larger cities. Second major driver behind urbanization trend is young Finns moving to cities in search of education or jobs and after re-building their social networks in these cities, never returning to their hometowns.

Urbanization as a phenomenon started relatively late in Finland, after the Second World War in the 1950s. However, the largest migration movements to the urban centers occurred in the 1960s and the 1970s. As the housing demand in urban areas quickly increased, it led to numerous construction projects of suburbs to accommodate for the increased housing demand.

[3] www.vero.fi

Nowadays, these suburbs and their concrete buildings form the majority of Finland's housing stock and they are in general perceived as lower quality buildings – majority of them are in great need of renovation or complete reconstruction in the foreseeable future. Apartments built after the largest migrant flows, i.e. from 1980s onwards, are focused in providing higher quality of life for the occupants and are hence often higher quality buildings.

The development of housing prices discussed in Section 2.1 shows the practical impact of urbanization in Finland's apartment prices. The apartment prices have clearly diverged between HMA and rest of Finland after tech bubble in 2000s when more and more population started to move to HMA in search of jobs or education. Huovari et al. (2002) discussed the impact of this phenomenon on the housing markets of rural areas. They discovered that the non-returners, e.g. education seekers, leave empty dwellings to areas and hence reduce the capital value of the overall markets apartments as supply of apartments increases and demand of apartments decreases. Furthermore, Huovari et al. (2002) argued that the increasing population in these urban centers is reflected to the housing market of the corresponding area with higher rents and apartment prices. Since 1950s and the beginning of urbanization trend, it has become a major driver in the apartment markets, and according to Statistics Finland the urbanization rate in Finland has reached 69 percent in 2015.

However, beyond the need and urge to live in urban centers, there are multiple other factors influencing Finnish consumer's living preferences. Juntto (2007) has conducted one the most comprehensive studies regarding Finnish consumers and their housing preferences. The study is based on survey results obtained from 3,455 personal interviews with Finnish households. The main discovery is that 86 percent of Finns would prefer owner-occupancy to rental dwelling. However, there is an interesting disparency when examining the percentage of respondent's owning an apartment, standing at 63 percent. Juntto's (2007) survey respondents also reported an increasing preference for larger houses. In the past few years, this has slowly started to turn into reality:  the average square meters per dweller in Helsinki area has increased slightly from 32.46 square meters in 2001 to 33.79 square meters per dweller[4].

When it comes to the rental market, her study revealed that over a third of tenants claim that their inability to purchase house or apartment, e.g. due to problems saving the principal amount for acquiring a new house or apartment, is their primary reason for rental-dwelling in the first place. Moreover, she discovered that small niche-groups form large majorities in the rental

[4] www.hri.fi

markets, e.g. 94 percent of students are rental dwellers. However, it is important to keep in mind that majority of students are accommodated by the publicly subsidized sector in student housing apartment. Juntto's findings imply that the rental market in Finland is largely populated by consumers who are unable to enter owner-occupancy due to e.g. lack of funding, and hence do not have the opportunity to choose freely where and how they live.

These discussed two trends, urbanization and lack of opportunity to choose between rental and owner-occupancy, are also highlighted as key drivers for consumers' decisions by studies focusing solely on HMA. For instance, Backman (2015) examined solo-dwellers in Helsinki area in a survey analysis conducted in 2012. The study revealed that solo-dwellers preferred more urban housing when compared to rest of the population in Helsinki. Solo-dwellers seemed to prefer areas in close vicinity to urban centers and services. Backman (2015) also discovered that the segment of solo-dwellers is a very diverse group where dwelling preferences can vary to a large degree due to e.g. age of the dweller.

Similar study on the HMA was conducted by Häkkänen (2015), who conducted an interview study of rental dwellers. The study discovered that the decision to rent instead of acquiring an apartment was largely driven by convenience or changing life situation. Highlighted examples of these situations were retirement and beginning of higher education studies. Backing Juntto's (2007) findings, the study's another noteworthy discovery was that most of HMA's rental dwellers see the arrangement as a temporary option and would prefer living as owner-occupants in the long-term. The paper also discovered that the decision of choosing apartment is driven largely by apartment's location and quality of the apartment. Furthermore, many survey participants explained that rental dwelling enabled them to live in areas out of their reach as owner-occupant's due to the price level of apartments in the preferred area. Tuominen's (2014) discovery that dwellers living in areas characterized with higher socioeconomic levels, e.g. higher income or lower crime rates, were more satisfied with their living conditions. This increased satisfaction from living in areas usually characterized by higher apartment prices could partly explain Juntto's (2007) findings, rental dwellers prefer renting apartments in expensive areas rather than acquiring an apartment from area, which could potentially decrease their dwelling satisfaction.

# 3. THEORETICAL BACKGROUND

In this chapter, I first review the key finding regarding studies on disaggregated apartment and rental apartment markets. Secondly, I briefly discuss the importance of rental yield and capitalization rate in terms of real estate investments. Finally, I discuss the empirical research regarding the determinants of net rental yield.

## 3.1 Disaggregated rental markets

Majority of the studies on rental markets focus on explaining the determinants of rental yield with aggregated data, combining all types of apartments from different areas to one sample. This is often rationalized with the limited availability of data on individual rental apartments. However, the existence of single aggregated rental market seems illogical. Anecdotal evidence often suggests that one-room apartments are optimal apartment investment vehicles due to their smaller size, the variation in absolute rents is often seen as smaller than the variation in prices when apartment size increases. Hence, the rental yield from these smaller one-room apartments is seen as superior to larger apartments. Furthermore, location is often perceived as a key determinant to higher rental yields from real estate investments (Michaels and Smith, 1990). Both examples indicate that by simply treating apartment rental markets as a single aggregated mass, the analysis misses on the nuances and relative importance of certain determinants, e.g. apartment’s type and location.

The academic literature conducted with disaggregated samples fortifies this hypothesis. The most common way to disaggregate apartments markets in academic literature is to divide the market according to geographical areas or the characteristics of the dwellings (Bourassa et al. 2003). According to Michaels and Smith (1990), the geographical areas are often chosen by natural landlines, official city district borders or compass directions.

One of the first academic papers examining the disaggregated apartment markets was published by Schnare and Struyk (1975). However, this paper focused on establishing the theory of segmented apartment markets rather than examining the differing rental or price functions within the segmented markets. Segmentation in the study was conducted based on whether the apartment was located in the inner or outer suburbs of Boston. The paper did not examine whether the price functions were statistically significantly different in terms of regression slopes or coefficients. Despite its’ shortcomings, this paper can be perceived as the foundation for studies examining different apartment submarkets. Another example of early studies examining simple disaggregated markets was conducted by Allen et al. (1995), who discovered that the determinants of absolute rent levels differ significantly between the aggregate sample and disaggregated samples between different property types. This study was more focused on assessing different regression models and aimed to specify the optimal hedonistic model to assess rental equations. However, it hints to the importance of treating the apartment market as a heterogeneous market, instead of a homogenous aggregated market.

Michaels and Smith (1990) took one step further and examined apartments markets on a disaggregated level in Boston area to test for statistically significant differences between apartment submarkets. They challenged the general perception that apartments in general have single price function within a city, and that the equilibrium price for housing assets can be estimated with this said equilibrium model, disregarding the geographical location of the building. They viewed the underlying assumption behind the hedonic framework - that all agents in the market are familiar with all information to evaluate their options - challenged when it comes to apartment markets, which consist of different geographical areas or apartment types. They found significant evidence regarding apartment market disaggregation based on their specification regarding the quality of the neighborhood; premier location, above average neighborhood, average neighborhood or below average neighborhoods.

After the first papers focused on disaggregated markets, there has been multiple academic papers contemplating whether the appropriate modeling of rental markets is done on aggregated or disaggregated level. For example, paper by Berry et al. (2003), studied the rental yield in Dublin's apartment rental markets with multiple hedonic models. In this paper, the authors discovered that modeling the apartment rental markets at the disaggregate level instead of the aggregate level yielded significant improvements in model estimation.

However, the academic literature dividing apartment rental markets in submarkets based on the type of apartment is very limited. One of the few studies on this field was conducted by Wolverton et al. (1999), who studied apartment rental market disaggregation by unit type in greater Seattle area. He studied the determinants of rental level variation in apartments with different compositions of rooms and bathrooms. The authors concluded that the greater Seattle apartment rental market is not homogenous, and the largest differences came from the size, time, location variables and property's age. Even as multiple studies report diverging rental determinants between apartment submarkets, studies conducted on aggregate samples are still conducted on a regular basis due to insufficient granular data to examine disaggregate apartment rental markets. For example, a recent paper by Hanink et al. (2010) examined the newly developed Chinese apartment market and due to insufficient granular data, the authors were only able to use aggregate level variables.

The concept of disaggregated apartment markets has also been used in studies going beyond the physical and location related attributes of the apartments, and within different real estate assets. For example, Berlemann and Freese (2013), examined the impact of monetary policies on apartment markets and found that positive interest rate shocks have adverse effect on house and condominium prices and rents but no impact on commercial property prices. As another example, Bischoff and Maennig (2010) discovered significant differences in rental levels between apartments leased by private, public and association landlords in Germany. Multiple disaggregated market studies are conducted also on real estate assets beyond traditional apartments. For example, Dunse and Jones (2011) studied the difference of rental yield within office submarkets in Glasgow city. The market segmentation was done based on real estate agents' views on the market. The paper identified significant variation within the rental yields between assets perceived to be of different quality.

In the context of Nordic markets, studies of disaggregated apartment markets are scarce. Janssen et al. (2001) studied the capitalization rates in Stockholm across building type, age of the building and four specific locations with Stockholm city. The paper was more focused on estimating the accuracy of different equilibrium models than analyzing the reasons for variations in capitalization rates between the geographical areas. The only study taking into account the apartment submarkets in Finland was conducted by Oinonen (2013), who studied segmented apartment markets based on geographical areas, when examining the average sale times of Finnish apartments. He found that that apartments in the Helsinki area have higher probability to be sold within the first week when compared to other areas.

The numerous academic papers and their empirical findings show that researchers should treat the rental market as disaggregated instead of treating it as a single aggregated market. When it comes to rental market disaggregation in the Nordics, there are no studies comparable to my study to the best of my knowledge. Hence, the purpose of this study is to fill this void in the academic literature and examine the disaggregation of rental market in Finland, focusing on the HMA.

### 3.2 Concept of rental yield

In this section, I first open the fundamental theory regarding rental yield to form a common context for the importance of yield in real estate investing. Later in this section, I discuss the key factors determining rental yield from real estate investments. Depending on the availability of data from HMA, I employ these key determinants in my research.

### 3.2.1 Relationship between property and rental markets

Real estate is in essence a long-lasting commodity and its' price and production quantities are determined in either asset or capital markets. Hence, to understand how the rent level of a real estate asset is formed, it is fundamental to understand the market for the usage of real estate. This market is often referred as the property market. Originally, the fundamental link between these markets was first defined by DiPasquale and Wheaton (1996), who argued the there are two links between the markets. The first link is between the rent levels in the market and asset prices in the market. The second link between is between the construction supply and asset prices in the market.

Following DiPasquale's and Wheaton's (1996) argument, when investors are purchasing a real estate asset they are in practice purchasing the right to cash flows from that asset. Appreciation in the cash flow hence increases the discounted cash flows from the asset, leading to asset price appreciation. Similar logic applies to the link between construction supply and apartment prices. If the construction supply of apartments increases in geographical area, hence increasing the supply of properties to the market, it hence drives down the prices of properties as the market balances in new equilibrium. Moreover, the rent levels in the markets are impacted by the preferences of rental-dwellers, driving real estate assets with preferable qualities to higher rental levels. Similarly, developers often undertake construction projects of apartments which have preferred qualities for either owner-occupants or rental-dwellers, depending on the purpose of the project.

This study focuses on the link between rental income and asset prices. This relationship is often defined as the rent-to-price ratio or capitalization rate in the real estate literature and practice. The equilibrium capitalization rate is impacted by the long-term interest rate of the economy, expected price appreciations and rent increases, risks associated in renting property and the taxation of property assets. However, in shorter time-periods, the capitalization rate might deviate from the long-term equilibrium capitalization rate.

Due to these variations from the long-term capitalization rate, it is imperative for rental investors to understand which determinants are driving the capitalization rate of their assets as it is directly linked to the returns from real estate assets. In the following section, I review the academic literature discussing the determinants impacting the rental yields, i.e. capitalization rates, of apartments.

### 3.2.2 Determinants of rental yield

Due to the importance of rental yield in the context of real estate investments, the academic literature on the subject is plentiful. Malpezzi (2003) wrote a review of the possible variables that can be used to predict rents or housing prices. In his study, the most important factors are the number of rooms, floor area, structure type, age of dwelling, neighborhood´s socio-economical characteristics, distance to central business district or sub-centers of employment or education, time of data collection and tenant characteristics if available. Similarly, paper by Sirmans and Benjamin (1991), synthesizes large number of studies regarding the determination of market rent. In Finnish context, Laakso (1997) investigated housing prices using transaction level data between the areas in Helsinki metropolitan area. This study was followed by e.g., Moilanen and Terho (2010), who built upon Laakso's study to examine the cross-sectional variation of net rental yields in Finnish housing market.

It is also noteworthy to note that often the rental yield determinants are universal from apartments to houses. Kain and Quigley (1970) conducted one of the first papers examining the determinants of rental yield in houses and discovered that physical attributes, e.g. the size of the apartment, the number of amenities and lot size were significant in explaining the house rents. Furthermore, quality related variables, based on a ranking from one to five, were highly significant in explaining the rental level of the house. To continue, Miller (1981), defined that the key determinants in house rental levels are physical conditions and the location of the apartment. Cross-referencing the list of rental yield determinants between houses and apartments implies that largely similar variables determine the apartment and house rents.

Synthesizing the past literature on rental yield determinants from both apartments and houses, the key categories determining the rental yield of apartments or houses are: a) apartments physical attributes, e.g. floor area and structure type, b) location related variables, e.g. vicinity to central business district, c) neighborhood related variables, e.g. the unemployment rate in a postal code area. Furthermore, the literature finds numerous other key determinants, not directly linked to the previous categories. For the purpose of this study, these determinants are labelled as d) other general characteristics, e.g. the vacancy rate of rental apartments. In the following, I review the academic research and findings related to these four categories in larger detail.

*a) Apartments characteristics*

In the Finnish context, Laakso (1997) examined the impact of apartment's structure type on the dwelling prices in HMA. He discovered that dwelling prices increase linearly with apartments size, with the exception of apartments larger than 200 square meters. Terho and Moilanen (2010), discovered that age of the apartment is highly significant attribute defining the net rental yield of apartments. In general, larger apartment sizes led to decreasing net rental yield and vice versa. Furthermore, they examined how net rental yield is affected in apartments built between 1960s and 1970s. As discussed in the previous section (please see Section 2.3), the apartment quality in Finland is viewed as subpar in apartments built in the 1960s and 1970s due to large development projects of low quality concrete buildings to accommodate migrant flows from rural areas to cities. Their results implied that the net rental yield was higher in these apartments due to the lower acquisition prices.

Similarly, Melakari (2014) discovered that the rentable area, age and location of a flat had the largest impact on apartments rental level. In addition, he found that new apartments commanded higher rental yield levels compared to older apartments in the southern parts of Helsinki. When it comes to other characteristics of apartments, the studies often incorporate dummy variables regarding the existence of e.g. elevator in the apartment building. The paper by Nikola (2011) and Moilanen and Terho (2010) found that elevator in the apartment increased the sales prices of the apartments. Furthermore, Nikola also studied the impact the condition of the apartment has on sales price. She discovered that apartments classified as being in good condition by the real estate agents were on average 15 percent more valuable than apartments classified as being in bad condition.

International studies largely concur on these rental yield determinants. For example, Sirmans and Benjamin (1991), discovered that the rental yield of apartments often deteriorates with the age of the apartments. However, in regards to the aforementioned importance of apartment structures, international literature is non-existent. Majority of international studies are conducted in US, where granular apartment level data is rare and majority of the academic studies are conducted on aggregated samples with limited apartment level information.

Due to the reported importance of variables related to physical attributes of apartments, I also use these in determining whether the rental markets in HMA are heterogeneous between one-, two- and three-room apartments. I use variables which have been employed in previous studies to ensure comparability to existing literature. I employ variables related to the size of the apartment, a variable for new apartments, variables representing the age of the building, i.e. apartments built between 1960s and 1970s and apartments built after this period, variables representing if the apartment is in the top or bottom floor and finally dummy variables for the existence of sauna, balcony, owned lot and elevator in the apartments.

*b) Location related variables*

Alonso (1964) formed the basis for modern microeconomics of urban economics by arguing that the location of apartments has a significant role in determining the demand, supply and value of apartments. He presented a model of household's location decision based on consumer theory. After hedonic pricing models were developed (see e.g., Lancaster, 1966; Rosen 1974), academics started to incorporate location related variables in determining the housing prices and rent levels. Following the establishment of location as the determinant in housing prices and rents, Benjamin and Sirmans (1994) examined the effect of mass transportation on apartment rents. They discovered that in Washington D.C rents fall, on average, 2.6 percent for every one-tenth mile decrease in the distance to the closest mass transportation station, in this case from a metro station.

In a similar fashion, multiple studies in Finnish context have adopted location related variables in their hedonic regressions explaining either apartment prices of rental levels. Laakso (1997) employed multiple location related variables in his extensive study of Finland's apartment markets. The most relevant factors in his study were the distance to the central business district (CBD) and the distance to closest railway and metro stations. Proximity to Helsinki CBD had high impact on the prices in both economical and statistical spectrums. This is no surprising as the apartment stock in close vicinity to city center is limited. In regards to railway or metro stations, Laakso's (1997) findings indicate that the positive impact of accessibility outweighs the possible negative externalities from the vicinity of rail or metro stations. However, when he leaves out the dummy variables controlling the general neighborhood, the coefficients for metro stations turn negative, implying that neighborhoods right next to metro stations can be perceived as negative for home acquirers.

To continue, also the newer studies in Finnish context reach similar conclusion about the proximity of transportation links has on apartment prices and rent levels. For example, Joutsiniemi (2011) examined the impact the announcement of metro line extension to west had on the apartment prices sold on that area. He discovered that apartments in close proximity to planned West Metro stations increased in value when compared to other apartments in Espoo region. Furthermore, Melakari (2014) found that apartments close to metro or train stations have lower rent levels than apartments farther away.

Finally, multiple studies also incorporate the proximity of university campuses or higher education centers to apartment price or rent equations. Ogur (1973) discovered that the presence of a college or university in a geographical area caused the price of housing services to be higher in comparison to other geographical areas. Moreover, Guntermann and Norbin (1987) examined the determinants explaining the variability between apartment rents. They discovered that the rental yield determinants differed between apartments close to university campuses to those further away. Furthermore, Sirmans and Benjamin (1991) reported that the preference of students to live closer to their campuses might have a positive impact on the areas rents.

To control the impact of the apartment's geographical location, I use variables expressing the proximity of the closest university campus, metro station or railway station from the apartment's geographical location.

*c) Neighborhood related variables*

The importance of neighborhood variables is derived from general theories regarding segregation. According to the theory, households prefer homogenous neighborhoods, i.e. neighborhoods with similar dwellers from similar socio-economical standing (See e.g. Fujita, 1989; Li and Brown, 1980). According to Laakso's (1997) literature review, almost all academic papers have found significant relationship between apartment prices and neighborhood's socio-economic status.

Furthermore, in his own study Laakso (1997) used multiple variables explaining the socio-economic status of the neighborhoods. These included the income level of the area, unemployment rate, size of households, population with foreign origin, share of owner-occupied dwellings and education level. However, he noted that only a few of these variables can be simultaneously employed due to multicollinearity issues. In the context of this study's scope, it is important to note that Laakso was specifically studying apartment prices. For example, Juntto (2007) discovered that the rental market in Finland exhibits a concentration of low-income households. Hence, the results for socio-economic factors might be different in the context of rental markets.

Due to data limitations, this study is not directly including socio-economic variables in the model, i.e. I do not include variables regarding the unemployment rate in the apartments neighborhood in my full models. However, I had access to the median income and unemployment rate on postal code level in HMA between the years 2012-2014. Hence, I test the importance of these variables in my robustness tests section (please see Section 6.5.4) to identify whether the neighborhood characteristics have significant impact on the net rental yield. In the main analysis, I control the postal code areas to take into account the importance of location.

*d) Other variables: Vacancy rate and supply of apartments*

Vacancy rate is important determinant in the net rental yield as theoretically the rental yield should be adjusted downwards for the vacancy rate of the apartment during the ownership period. As an example, Garner and Verbrugge (2009) adjusted their predicted rental levels downwards with the equivalence of region wide vacancy rate. Moreover, Hagen and Hansen (2010), in their study of the natural vacancy rate in Seattle Metropolitan area, discovered that the natural vacancy rate in the market varied between years. Furthermore, they also discovered small variation in vacancy rates between apartment types. Similar findings were done by Glaeser and Gyourko (2007), who noted that there is potential variance in the vacancy rate between the expected tenure of different types of households. This would lead to variation in the net rental yields depending on the tenant profile.

However, due to insufficient data of HMA's vacancy rates, I have to leave vacancy rates out of the scope of my study. However, as my study focuses solely on one small geographical area, large variations in the vacancy rate are unlikely. Only large variations in the vacancy rates within of one-room, two-room or three-room apartment groups within HMA would cause biases in my results. I assume that the vacancy rate is rather steady in HMA within different apartment groups, and thus the uncontrolled vacancy rate does not distort my empirical results.

Another noteworthy variable explaining rental levels of apartments is the supply of new rental apartments in the market. Chichernea et al. (2008) examined the impact of supply side variables, e.g. available capital, development regulation and new constructions have on apartment prices or aggregate rent levels. They discovered robust evidence that the supply constraints have suppressing impact on the capitalization rates in property markets. Furthermore, Hanink et al. (2010) studied the China's apartment rental markets on aggregate level. In their study, they assessed the rental market on both local and nation-wide level. The locally specified model found that the spatial variables, especially the supply of new apartments available at market prices influenced the rent levels positively in the same geographical area.

To incorporate the importance of apartment supply in my study, I include the supply of new apartments on the postal code area to my model. The variable represents the apartment supply during the past two years in that postal code area and my study is the first, to the best of my knowledge, exploring the impact of apartment supply on rental yields.

# 4. METHODOLOGY & HYPOTHESES

In this chapter, I discuss the methodology employed in the study and my key hypotheses. Firstly, I explain the methodology used to examine the determinants of net rental yield in previous academic studies. Secondly, I focus on the specific methodology to test the hypothesized disaggregation of the rental markets. I discuss multiple different statistical tests aimed to examine the statistically significant differences between the rental submarkets. Finally, I present the testable hypotheses of my study.

## 4.1 Hedonic equation

Basic hedonic theory suggests that the market value of a commodity product is formed by its' characteristics (Rosen, 1974). As the name implies, hedonic functions analyze the demand and prices for varying sources of satisfaction and gratification. In the context of rental market, this sums up to all the attributes that constitute to the fair rent level of a dwelling. The fundamental equation, as illustrated by Malpezzi (2003), is the following:

$$R = f(S, N, L, C, T)$$

where R = rent, S = structural characteristics, N = neighborhood characteristics, L = location within the market, C = contract conditions or characteristics and T = the time the of rental or sales transaction is observed. Similar equation is applicable to different rental market subgroups, in which case the model can be defined separately for these different subgroups.

Within real-estate literature, there exists multiple different specifications of the hedonic equation and large debate regarding the optimal model specification. Halvorsen and Pollakowski (1981) proposed a functional form including two different Box-Cox transformations and interactions terms. Cassel and Mendelsohn (1985) built on these findings, stating that choosing the optimal function is in effect a trade-off depending on the use of the function. They stated that often the model with the best fit is the most appropriate model, but the results might be suboptimal if the relative importance of the different attributes is the main interest. Moreover, Cropper et al. (1988) conducted a throughout model selection investigation by simulating linear, log-linear- quadratic, linear Box-Cox and quadratic Box-Cox models. They concluded that Box-cox linear form had the smallest average bias between the models, but the linear form produced the smallest maximum bias in the simulations.

In Finnish context, Terho and Moilanen (2010) conducted similar examination of optimal model specifications in Finnish rental markets. They examined different Box-Cox specifications and linear models on a dataset from HMA. After reviewing the model results, they ended employing double-log specification in their study. Similar conclusions were reached by Arimah (1997) and Laakso (1992) who employed double-log specification due to factors related to model interpretation, significance and stability. Similarly to Terho and Moilanen (2010), Laakso (1997) tested Box-Cox transformation and deemed that “the results did not change the main conclusions based on the results of log linear models”.

Due to the results from previous studies and the fact that Terho and Moilanen (2010) recently specified their model accuracy specifically with HMA data, I chose to employ double-log specification in my empirical analysis. Hence, instead of contemplating on the optimal model specification, I rely on the robust results from existing empirical studies, and focus my study on examining the rental market disaggregation. In other words, in the first part of my empirical analysis, I employ double-log regression specification to examine the differences in rental yield determinants between aggregated sample regression and apartment subgroup, i.e. one-, two- and three-room, regressions.

### 4.2 Examining disaggregated rental markets

In this section, I explain the exact steps to assess whether the net rental yield determinants statistically significantly differ between one-, two- and three-room apartments. This process has two key steps. Firstly, I examine the statistical difference between the subgroup regressions, i.e. examine significant differences between the slopes of subgroup regressions. This is done with a Chow-test. Secondly, I examine statistical difference between the coefficients of the submarket regressions with Tiao-Goldberger F-test. This will reveal which of the coefficients are causing the statistical differences between subgroup regression slopes.

#### 4.2.1 Chow-test

After establishing the distinct double-log regressions models for different submarkets, I examine whether these distinct regressions are statistically significantly different from each other. Wolverton et al. (1999), employed similar methodology in their study to examine differing rental yield determinants between apartment types. Similar methodology has also been employed in multiple studies focusing on distinguishing apartment submarkets based on geographical areas, e.g. Sherif's and Ashraf's (2007) study of Cairo's apartment markets and Berry et al.'s (2003) study of Dublin's apartment rental markets.

This examination is conducted by a series of Chow-tests (1960) to formally examine the null hypotheses:

$$\underline{\beta}_{1-room\ apartments} = \underline{\beta}_{2-room\ apartments}$$

$$\underline{\beta}_{1-room\ apartments} = \underline{\beta}_{3-room\ apartments}$$

$$\underline{\beta}_{2-room\ apartments} = \underline{\beta}_{3-room\ apartments}$$

To test these hypotheses, I run distinct aggregate models for the combinations of one- and two-room apartments, one- and three-room apartments and two- and three-room apartments. In addition, I run regressions based on the apartment type subgroups within these aggregate samples to compute the Chow-test statistic. The formula to perform the F-distributed Chow-test is the following:

$$F = \frac{SSR_F - (SSR_{S1} + SSR_{S2})}{(SSR_{S1} + SSR_{S2})} x \frac{(N_{S1} + N_{S2} - 2k)}{k}$$

where:

$SSR_F$ = Sum of squared residuals from the aggregated regressions

$SSR_{S1}$ and $SSR_{S1}$ = Sum of squared residuals from segregated regression 1 and 2

$N_{S1}$ and $N_{S2}$ = total number of observations in segregated regressions 1 and 2

$k$ = degrees of freedom in the regressions, equaling the number of parameters plus 1

### 4.2.2 Tiao-Goldberger F-test

As a second step in analyzing the possible existence of disaggregated rental markets in HMA, the dissimilarity between independent variables is examined by employing Tiao-Goldberger (1962) test for differences between the regressions coefficient estimates. Similar methodology in validating the existence of submarkets have been used in the previous academic literature by e.g. Wolverton et al. (1999), Black et al. (1997), Allen et al. (1995), Michaels and Smith (1990) and Fuerst (2008).

The null hypothesis for Tiao-Goldberger F-test is the following:

$$\beta_{i(1-room\ apartments)} = \beta_{i(2-room\ apartments)} = \beta_{i(3-room\ apartments)}$$

Where coefficient i= 1 to k. The Tiao-Goldberger test is F-distributed with $((L-1, N_{1-room} + N_{2-room} + N_{3-room} - Lk$ degrees of freedom. The formula to calculate the F-statistics is the following:

$$F_{TG} = \frac{\sum_{j=1}^{L} \frac{\left(\hat{b}_{ij} - \bar{b}_{ij}\right)^2}{P_{ij}}}{\sum_{j=1}^{L} SSE_j} x \frac{\sum_{j=1}^{L}(T_j - k_j)}{L-1}$$

Where:

$$\bar{b} = \frac{\sum_{j=1}^{L} \frac{\hat{b}_{ij}}{P_{ij}}}{\sum_{j=1}^{L} \frac{1}{P_{ij}}}$$

And where:

L = number of models being compared (three in this case)

$T_{ij}$= number of observations in model *j*

$k_j$= number of variables in the model *j*, including the intercept

$\hat{b}_{ij}$= the OLS coefficient estimator for parameter *i* in model *j*

$P_{ij}$= the diagonal element of the *i*th parameter of $(X'X)_j^{-1}$

The Tiao-Goldberger test compares each estimate on the weighted sum of parameter estimates across all models (Michaels and Smith, 1990). The resulting F-statistic reveals which of the coefficients vary significantly from one another across the one-room, two-room and three-room subgroups.

**4.3 Key hypotheses**

In this section, I outline the main hypotheses of this study. The fundamental hypothesis is to examine whether rental markets in HMA are disaggregated. In other words, is the net rental yield defined by different coefficients in different apartment types, more specifically between one-, two- and three-room apartments. The hypothesis is the following:

**H$_1$:** There are different rental micro markets in different types of apartments, namely one-, two- and three-room apartments in HMA – the net rental yield is formed by different apartment characteristics in these different submarkets

To examine hypothesis number one, I firstly construct distinct double-log regressions for the aggregated sample, containing all types of apartments and distinct regressions for the subgroups of one-, two-, and three-room apartments. I examine whether and how the statistical significances, signs and magnitudes of variables vary between these regressions. Secondly, I examine whether these regressions are statistically different from another with the aforementioned Chow-test (Chow, 1960). Finally, I employ a Tiao-Goldberger test to examine which of the coefficients between the subgroups are statistically significantly different between the regressions.

Furthermore, as this study is the first academic paper incorporating supply side variables to explain net rental yield, i.e. the new apartment supply in cubic meters for the two preceding years, my second hypothesis is to examine how this impact's the net rental yield of one-, two- and three-room apartments. This hypothesis is aligned with previous studies regarding apartment supply (see e.g. Chichernea et al., 2008), which have discovered that the supply of new apartments in the area suppresses the rental yield levels within the same geographic area. Hence, the hypothesis is stated as:

**H$_2$:** High supply of new apartments preceding the apartment rental advertisement suppresses the apartment's net rental yield

My third hypothesis is built on the foundation of previous literature on Finnish rental markets, e.g. Moilanen and Terho (2010), who discovered that net rental yield is negatively impacted by the apartment's size and the function between apartment size and rental yield is concave as the importance of size diminishes in larger apartments. I hypothesize that this impact is largely driven by one-room and two-room apartments, which are characterized by higher prices per square meter and customer segment consisting largely of lower income customers, e.g. students. Thus, I hypothesize that in these apartment submarkets, larger acquisition prices from larger space is not compensated in terms of higher net rental yield as the customer group has limited disposable income on housing. This hypothesis is tested firstly by comparing the significance of size between the subgroup regressions and secondly by examining the impact the apartments size, type of apartments and the relation of size and type of apartments has on net rental yield. The hypothesis 3 is stated as:

**$H_3$:** Larger apartment size has adverse impact to the net rental yield only in the subgroups of one-room and two-room apartments

# 5. DATA DESCRIPTION

In this chapter, I explain the sources from which I have received my data and provide descriptive statistics of my sample. Firstly, I go through the main data sources in my study, Oikotie and KVKL databases, and the matching process to merge these datasets. Secondly, I discuss other key sources of data employed in this study. Finally, I present the descriptive statistics of the data and discuss the potential limitations of my dataset.

## 5.1 Dwelling specific variables

My main sources of dwelling specific data are Oikotie rental advertisements and KVKL transaction data. Oikotie is Finland's largest apartment rental and sales website and a subsidiary to Sanoma Oyj, which is listed to OMX Helsinki. KVKL is Finland's association for real estate agents and their price database contains apartment and transaction characteristics, including realized transaction prices of all transactions orchestrated by real estate agents in Finland from the year 1998. I use only observations from HMA in my sample to ensure highest possible quality in my sample. The limited scope allows me to firstly manually verify the correctness of the data and ensure robust matching of rental advertisements to realized apartment transaction to estimate the true net rental yield. Secondly, multiple sufficiently granularly collected independent variables were only available in HMA.

### 5.1.1 Oikotie apartment advertisements

I received data from Oikotie containing all Finland's rental advertisements from 2012 to 2016. Sample of HMA region rental advertisements contained 112,331 transactions. Out of these, 60 percent of transactions were in Helsinki and rest in Espoo and Vantaa region.

In addition to the asked rental price, the dataset contains multiple other descriptive characteristics. It contains the full address of the apartment, the postal code, district and city, publication and expiration date, floor of the apartment, construction date, balcony, floor area in square meters, condition of the apartment and type of apartment to name of few. Unfortunately, I did not gain access to descriptions written by the selling part to include this qualitative information in my empirical analysis.

### 5.1.2 KVKL's apartment transaction data

The data from KVKL's database was hand collected. Observations range from the start of January 2012 until the end of December 2016 in HMA. In total, I retrieved 85.926 distinct real estate transactions.

The data contains similar apartment data as the Oikotie apartment advertisements dataset, but the quality of the data is higher. For example, the data does not have as many missing values from key characteristics and spelling mistakes are rarer. The plausible explanation for this is that the data provided by Oikotie is subject to larger human error when submitting new rental advertisements. Most likely, the inputs in KVKL database are examined more carefully or the absence of transaction without real estate agents improves the quality of the data. The KVKL dataset contains following transaction specific information: type of the dwelling, dwelling address, postal code and district, size of dwelling, building year, amount of rooms, total number of floors in the apartment and floor of the dwelling, transaction price, portion and debt and debt-free portion, price per square meter, whether the apartment is a new, condition of apartment, maintenance charges and various information about the timing of the transaction.

The procedure to clean and match these two distinct datasets is explained in the next section. To my best understanding, this study is the first one to match actual transaction prices to advertised rental prices. Hence, creating a novel dataset with accurate net rental yield information.

### 5.1.3 Matching process

The matching process is performed with cleaned databases, Oikotie database having 103,788 observations and KVKL having 63,138 observations. The cleaning process comprised of removing other than apartment rental and sales information, removing observations with insufficient information, e.g. missing rental or transaction price and removing defective observations, e.g. rental advertisements with spurious addresses.

As a last step, I manually ensured the robustness of all of the matching criteria, e.g. the addresses of the apartments, by manually correcting possible spelling errors and finding and adjoining missing information to the observations from both data sources. This process consisted of hand-collecting missing address information, e.g. finding postal code and district information for observations with address information but no postal code and manually correcting numerous address spelling mistakes from both Oikotie and KVKL databases. This process was done to ensure the highest possible quality of data in sample and because both datasets contained multiple observations with spelling mistakes in the address fields, which would have reduced the accuracy of matching process and decreased my sample size.

The matching procedure of Oikotie’s and KVKL’s data was completed with the following criteria. Firstly, the full addresses of the apartments had to exactly match each other. Secondly, postal codes of the apartments had to match exactly. Thirdly, the floor number had to be an exact match. Fourthly, the amount of rooms had to match exactly. Fifthly, the apartment’s space in square meters had to be within 10 percent logarithmic deviation from each other. The interval is rationalized to ensure that apartments even with misspelled size in either of the databases would be paired. Sixthly, the construction years of the buildings had to exactly match each other. Finally, the construction years of the buildings had to exactly match each other. Table 1 summarizes all of the matching criteria, the match type, possible matching baskets and the description of the criteria.

**Table 1 - Matching criteria**

This table summarizes the criteria used to match the two distinct datasets from Oikotie, containing the rental advertisement data, and KVKL, containing apartment transaction data. In total 7 criteria were used in the matching procedure. These criteria were the full address of the apartment, postal code, floor number, amount of rooms, apartment size, construction year and the transaction date. Of these 7 criteria, 5 were exact and 2 were range criteria.

| **Criteria** | **Match type** | **Basket 1** | **Basket 2** | **Basket 3** | **Description** |
|---|---|---|---|---|---|
| Full address of the apartment | Exact | | | | Full address containing the street name, street number, possible stair alphabets and apartment number |
| Postal code | Exact | | | | Postal code of the apartment |
| Floor number | Exact | | | | Floor of the apartment |
| Amount of rooms | Exact | One-room | Two-rooms | Three-rooms | Number of rooms in the apartment |
| Apartment size in square meters | Range | | | | Sizes within 10 percent logarithmic deviation |
| Construction year | Exact | | | | Construction year of the apartment |
| Transaction date | Range | | | | Rental advertisement can only be published after sales transaction |

This procedure yielded me a sample of 8,303 observations with exactly matching criteria. This matching procedure gave me the possibility to analyze the real net rental yield of the apartments based on actual transaction prices, instead of comparing ask prices to ask rents. Ask prices in real estate often do not give a full picture, as there is often room for negotiation. However, I am still employing simply the ask rents from rental advertisements instead of actual agreed rental levels. Despite this, rents are often set to meet investors yield expectations and I suspect there is only minimal room for negotiations. Hence, I argue that the advertised rental levels are fair representation of real rent levels.

Furthermore, as my study examines the net rental yield of apartments, I had to make two adjustments to the actualized transaction price and asked rent levels to compute the net rental yield. Firstly, I adjusted the transaction prices to include the transaction tax (2 percent in apartment houses in Finland). Secondly, I deducted the maintenance charges from the gross rent to operate with net rent levels. As the used transaction prices were full transaction prices, i.e. price with household debt and equity portions, there is no housing company loan payments to take into account when computing the net rental yield of the apartments.

### 5.2 Analysis of outlier observations

As the final step in constructing the dataset, a final manual outlier analysis was conducted to ensure the quality of the sample. The removed observations can be classified into either observations with insufficient or incomprehensive values and outlier observations.

Firstly, three observations being the only transaction in the specific postal code area were removed. Secondly, 380 observations with no maintenance charges were removed, because it was integral in calculating the net rental yield of the apartments. Thirdly, I removed 6 observations with no indication of rental time and 47 observations with a rental advertisement time over one year. I hypothesize that these rental advertisements never realized into a rental agreement. Finally, I removed 259 observations where the classification of apartment’s condition had risen by two or more steps on a scale of satisfactory, tolerable, good and excellent from the purchase transactions condition to rental apartments condition. These apartments have likely undergone financially significant renovations. By definition, the price of future renovations should be added and discounted in the apartments purchase price when calculating the net rental yield. Apartments which have undergone large renovations, would thus show upwards biased net rental yield as the cost of renovations is not calculated in the acquisition cost of the asset. In total, I removed 695 observations with insufficient or incomprehensive values.

The second subgroup of outlier observations consisted of observations with significantly higher or lower transaction prices or rental levels when compared to the postal areas average price or rental levels. These rent or price levels were classified as being more than 2.5 standard deviations from the mean of the specific postal code area. In total, I removed 145 observations with disproportionate price or rent levels. Moreover, I manually erased 12 observations with disproportionate maintenance charges. After looking at the observations more closely, I discovered that these were in fact advertisement of single rooms instead of whole apartments, and hence the maintenance charges did not reflect the actual costs. After removing the variables with insufficient or incomprehensible values and outlier values, the final sample consisted of 7.451 observations.

### 5.3 Other independent variables

In addition to the apartment specific variables from Oikotie and KVKL, I sourced variables related to the geographical location of the apartments, e.g. distance to closest public transportation hub, variables reflecting socio-economic factors in the postal code area, e.g. median income and unemployment rate, and variables explaining the supply of new apartments in the area. In the following, I will open these data sources to larger extent.

#### 5.3.1 Postal code related variables

Postal code level data regarding the socio-economic factors was retrieved from Paavo-database[5] from Statistics Finland. Statistics Finland is maintained and updated by the Finnish public authority, and it produces official statistics from Finland. From this database, I retrieved town and postal code data regarding the median incomes and unemployment rates. This information was matched to observations based on postal codes.

Due to the limited nature of Paavo-database, I was able to retrieve data regarding socio-economic factors only for the years 2012-2014. Hence, I am using these numbers only in my robustness tests, (please see Section 6.5.4). This allows me to keep the sample as large as in my main analysis regarding market disaggregation, but allows me to still examine the impact these socio-economic factors may have on net rental yield determination in HMA.

#### 5.3.2 Supply side variables

As discussed, multiple international studies, e.g. by Chichernea et al. (2008) have examined the impact of supply side variables, e.g. available capital, development regulation and new

[5] http://www.stat.fi/tup/paavo/index.html

constructions have on apartment prices or rents. To my understanding, this study is the first one on Finnish markets using supply side variables along with dwelling, geographic and spatial variables to analyze its' impact on apartment rent levels.

I received a unique dataset regarding the apartment supply in HMA from the Building and Dwelling Production department of Statistics Finland. The data contains information regarding the amount of newly finished residential buildings in cubic meters and number of built apartments. I selected to use the cubic meters as a proxy for apartment supply as it is, according to Statistics Finland, their main variable explaining apartment supply and hence has the highest data quality.

I used this information as a proxy for the supply of new apartments into the HMA rental market. The employed variable is the aggregate amount of finished apartment in cubic meters in the postal code area. The selected timeframe is two years prior to rental year to account for large variation in construction levels between years. Operating with only apartment supply from previous year would have had the following issues. Firstly, there is large variation in constructed apartments between the years. Only counting the previous year's supply could hide information about large finished construction projects in the near-term past. Furthermore, most rental contracts in Finland are long-term. Thus, the impact of only one-year's supply would probably have negligible effect on the areas rental level as most of the agreed rental contracts would remain during that time, hence failing to capture the possible implications to areas rents due to new developments.

Variables regarding the postal code areas, i.e. socio-economic factors and apartment supply, were then matched to my dataset based on the postal codes of the observations.

### 5.3.3 Other dwelling specific data Google Maps and HRTA

As discussed previously (please see section 3.2.2), geographical information has been found significant in studies examining net rental yield or apartment price development. For example, Benjamin and Sirmans (1994), discovered that rent levels were significantly higher for apartments close to mass transportation stops. Similarly, in Finnish context, Laakso (1997) discovered that apartments are often more valuable closer to metro and railway stations. Hence, I also incorporated variables explaining the apartments location in relation to key public services. In this section, I explain the process of obtaining the geolocations of my observations and key public services and the process of calculating distance between the apartment observations in my sample and public services, i.e. university campuses and metro and railway stations.

I retrieved the geolocations of the apartments in my sample from Google Maps[6] service. Oikotie provided few apartment coordinates, but multiple missing values led me to use Google Maps service to obtain coordinates for full sample. Furthermore, to receive the exact coordinates of railway and metro stations, I used the HSLs (Helsinki Regional Transport Authority) data archives. HSL is the state's transportation service providing all public transportation in HMA, including busses, metros, trams, trains and lorries.

Other variables obtained from Google Maps were the geolocations of major university campuses in HMA. In total, I retrieved the coordinates of **7** university campuses in HMA. From HSLs databases, I retrieved the exact coordinates of metro and railway stations. In total, I retrieved the exact coordinates of 18 metro stations and 41 railway stations. These stations included only operational metro stations and railway stations.

After retrieving the coordinates of my samples observations, university campuses and metro and railway stations, I calculated the shortest distance from each of the apartment observations to the closest university campus, metro station or railway station. I employed the Haversine formulae to compute the distance between these two coordinate points. The distance between the locations is always represented in kilometers. In practice, Haversine formula calculates the distance between any two points on a sphere.

$$haversin\left(\frac{d}{R}\right) = haversin(\varphi_1 - \varphi_2) + \cos(\varphi_1)\cos(\varphi_2)\, haversin(\Delta\lambda)$$

Where:

Haversin: Haversin function

D = distance between two locations on the circle of the sphere

R = sphere's radius

$\varphi_1$= latitude of first point

$\varphi_2$= latitude of second point

[6] Website accessible here: https://www.google.fi/maps

$\Delta\lambda$ =longitude separation

### 5.4 Variable overview

Table 2 below displays all of the independent variables used in my regressions. It displays variable names and descriptions.

**Table 2 - Variable summary**

This table summarizes all the variables used as independent variables in my study. The left column shows the variable names. The right column offers brief description of these independent variables. Last two variables are control variables.

| **Variable name** | **Variable description** |
|---|---|
| *Apartment size* | The size of the apartment in square meters |
| *New apartment* | Apartments built and rented during the same year |
| *Built 1962-1983* | Apartments built between 1962 and 1983 |
| *Built 1984-2016* | Apartments built between 1984 and 2016 |
| *Highest floor* | Apartment in the highest floor of the building |
| *Bottom floor* | Apartment in the bottom floor of the building |
| *Sauna* | Apartment with personal sauna |
| *Balcony* | Apartment with balcony |
| *Owned lot* | Apartment buildings situated on owned lots instead of rented lots |
| *Elevator* | Apartment building has elevator |
| *Construction supply* | Apartments completed on the postal code area during the previous 2 years in cubic meters |
| *University* | Distance to closest University campus in kilometers |
| *Metro* | Distance to closest Metro station in kilometers |
| *Railway* | Distance to closest Railway station in kilometers |
| *Rent year (control)* | Year the dwelling is rented to control for inflation in rental prices |
| *Postal code area (control)* | Postal code area of the dwelling to control for areal information, e.g. unemployment in the area |

### 5.5 Time and geographic distribution of the sample

In this section, I present briefly the descriptive statistics of the sample in terms of yearly and geographical distribution. Figure 3 displays the yearly distribution of the sample and Figure 4 the geographical distribution of the sample.

Figure 3 showcases how the sample is rather evenly distributed from 2012 to 2016. The year 2014 was the most active year in my sample in terms of volume and value. Both the value and the volume of the transactions starts to decline after this peak year and in my sample, the year 2016 has the lowest amount of transactions. One explanation for this is that the apartments sold in the end of 2016 are likely rented during 2017 and they are thus not visible in my sample.

**Figure 3 - Yearly distribution of transactions**

This figure showcases the number of sales transactions in my data sample and the value of these transactions in millions of euros. The figure illustrates how the number and volume has been rather stable from 2012 to 2016 in my sample. During the final years there is a slight decrease in both number and volume of apartment transactions. One likely cause is that often apartments are rented a while after they are actually sold, due to e.g. agreements about move in date. In my sample, this reduces the number of apartments sold in 2016 as they are not rented during 2016.

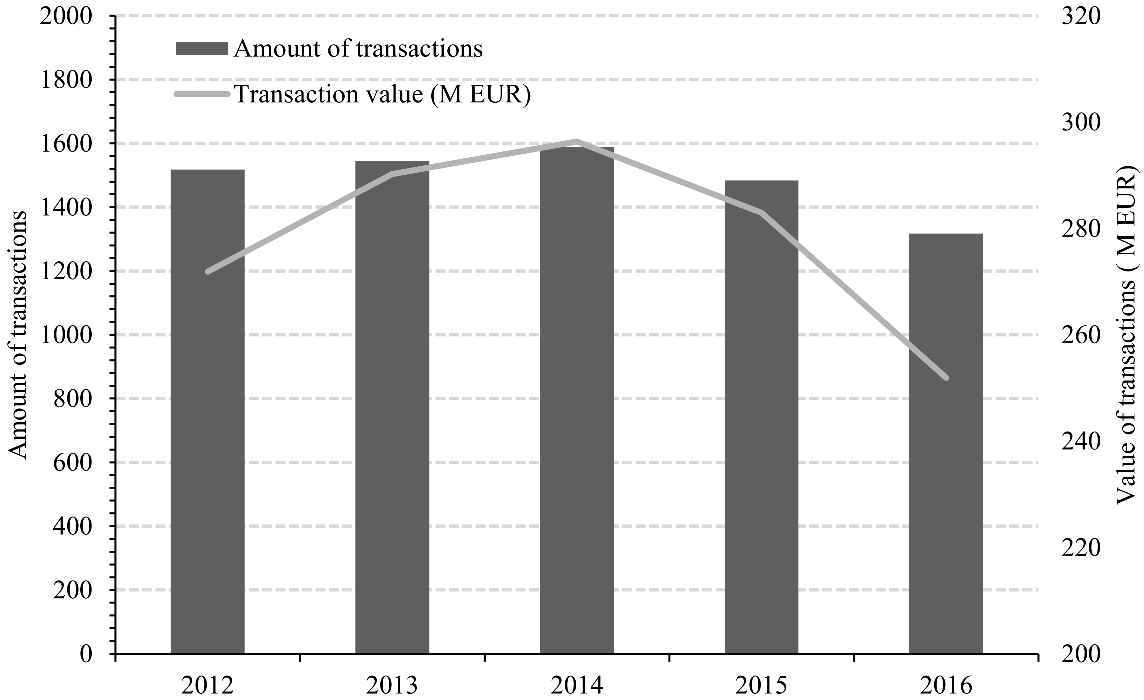
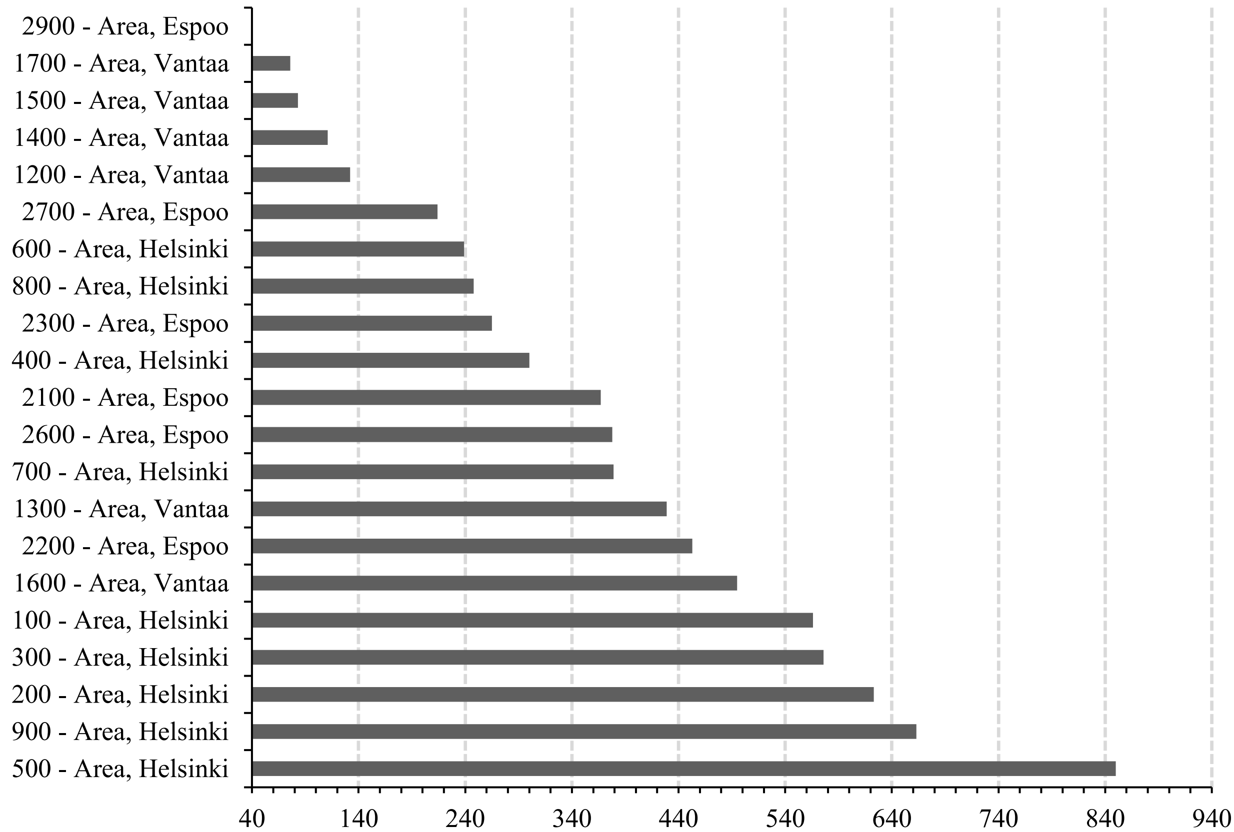



Figure 4 showcases the geographical distribution of the transactions based on postal code areas. Finland´s postal code system is based in standard international 5 number codes. First two numbers dictate the area and the last three numbers dictate the postal area. The areas are selected based on the first meaningful numbers, e.g. 100 – area contains all areas starting with 001. The selected criteria do not represent greater districts in Espoo and Vantaa area, but the areas are nevertheless geographically connected to each other and thus they have similar characteristics. Helsinki has the largest number of observation with a 68 percent proportion of total observations. Within Helsinki region, the most active area is Kallio. Within Espoo region, areas in close vicinity to Helsinki, such as Tapiola and Leppävaara, have the most observations in my sample. These regions are close to transportation links to both Helsinki centre and they are close to services. Within Vantaa region, growing and constantly developed Tikkurila area represents most observations in the sample.

**Figure 4 - Geographical distribution of transactions**

This figure showcases all of the transactions in the sample based on postal code areas. Timeframe ranges from 2012 to 2016. Finland´s postal code system is based in standard international 5 number codes. First two numbers dictate the area and the last three numbers dictate the postal area. The postal code areas are selected based on the first meaningful numbers, e.g. 100 – area contains all areas starting with 001. Helsinki is clearly the most active rental market in the sample, Kallio representing the most active area in my sample.




### 5.6 Potential data limitations

In this section, I discuss the main limitations and potential biases my dataset poses for my analysis. The major limitations of my dataset are the lack of distinctions between transactions from individual and institutional investors, missing information regarding vacancy rates, lack of transactions conducted without real estate agents, the lack of information regarding renovation projects completed or upcoming in the apartments and the non-exhaustive nature of my rental advertisements source. These limitations are explained and discussed in detail in the following sections.

#### 5.6.1 Distinction between individual and institutional investors

First limitation is that neither of the datasets distinguishes between retail and institutional investors. One could hypothesize that units advertised by professional real estate investors would be in better condition, located in prime-locations and have optimal rental apartment layouts. Hence, they would be listed on a rent premium when compared to the non-professional investors. This view is supported by the study of Sirmans and Benjamin (1991) who found that professional property managers have significantly positive impact on rents.

### 5.6.2 Vacancy rates

As discussed previously (please see section 5.2.2), vacancy rates have been found to have a significant effect on apartment rents. However, the results regarding vacancy rates have been at times conflicting. Thus, having access to market-specific vacancy data could have helped me to assess the impact vacancy rates have on ask rents. From theoretical standpoint, the vacancy rates should be embedded into the ask rents. However, in this context of limited geographical and apartment type scope, only large variance of vacancy rates within the HMA and within the specific apartment types, i.e. one-, two- and three-room apartments, would cause significant bias on results. Hence, I have not tried to speculate on vacancy rates and assumed that there is no significant variation in the vacancy rates within HMA.

### 5.6.3 No transactions by private sellers

The KVKL dataset contains only transactions made by real estate agents. This might cause bias in the dataset, as it does not contain real estate transactions brokered by home owners. The significance is likely very small as majority of Finns sell their apartments through real estate agents instead of pursuing sales on their own. However, this likely reduces the size of the dataset slightly.

### 5.6.4 Lack of information regarding large upcoming repair projects

One caveat in my dataset is that it does not include information regarding the completed or upcoming renovation projects to the apartments. These endeavors are often expensive projects for the rental investors. These projects are financed with housing company loans, which allocated on a pro rata basis based on the surface area of their apartments in relation to the total surface area of the building. For example, Nikola (2011) found that owner-occupant buyers do not count the price of the incoming pipe repairs correctly and often end up paying a premium when purchasing apartments with upcoming pipe repairs. However, the impact of renovations is indirectly taken in to the empirical analysis for the following reason. I am using the full transaction price when calculating the net rental yield. In other words, I count both equity and outstanding housing company loans as the purchase price. This treatment makes the apartments comparable in terms of past renovations. However, as I do not have information regarding the upcoming renovation projects, I am unable to include their impact in my empirical analysis.

### 5.6.5 Non-exhaustive list of rental listings

The main sources for the rental advertisements was the database of Oikotie, a subsidiary of Sanoma Oyj. However, this is not the only source of rental advertisements in Finland. The second largest service provider is called Etuovi, which is part of Alma Media Group. In practice, these two advertisement pages control the Finnish rental advertisement market. I hypothesize that even though it would be rational in terms of maximal exposure for renter to list the rental apartment on both portals, it probably does not happen with every available rental apartment. Reasons can range from brand preferences to the required effort to list on both pages, even as both allow free basic advertisements. Hence, even though most of advertises list their rental apartments on both forums, my dataset most likely does not contain all rental advertisements in HMA between 2012 and 2016.

## 6. RESULTS

In this chapter, I go through my empirical findings. Firstly, I go through the descriptive statistics to identify differing characteristics between the apartment types. Secondly, I examine whether the net rental yield determinants have varying statistical significances between the aggregated and disaggregated samples of one-, two- and three-room apartments. Thirdly, after establishing differences between the apartment subgroups in net rental yield determinants, I examine the statistical difference between the models and the individual coefficients. This part is conducted with a Chow-test and Tiao-Goldberger test. Fourthly, I examine the relationship between apartments rental yield, rent and price. Finally, I conclude this chapter by examining the robustness of my empirical results.

### 6.1 Descriptive statistics

In this section, I showcase the descriptive statistics of my sample. Table 3.1 below summarizes the mean, standard deviation, minimum and maximum values for all non-dummy variables. Similarly, Table 3.2 contains a frequency tabulation for the dummy variables. Both tables are arranged according to the apartment types.

Table 3.1. reveals that the sample is logically distributed when it comes to apartment rents, transaction prices and sizes between apartment types. Means, standard deviations and minimum and maximum values increase when transitioning from generally smaller one-room apartments

to larger three-room apartments. Furthermore, when it comes to distances to university campuses, metro stations or railway stations, the observations seem to be evenly distributed and the minimum and maximum values are similar. This implies that the sample is not biased with inconsistent observations.

Table 3.2. summarize the key information regarding dummy variables and reveals several characteristics regarding my dataset. Firstly, we can see that the number of new apartments is small in all subgroups. Most likely these are more often purchased for owner-occupancy than for rental purposes in HMA, due to the relatively high price of new developments. When it comes to the age of apartments, most of the one-room apartments are built before 1960s and most of the two-and three-room apartments are built between 1961 and 1983. This is aligned with the large urbanization trend starting in 1960s. In the efforts to accommodate the migrating population, most of the development projects have been made to accommodate larger family sizes instead of occupying only single-dwellers.

**Table 3 - Descriptive statistics of the sample by apartment type**

This table showcases the summary statistics for all independent variables except control variables used in the empirical analysis. Table 3.1 showcases the descriptive statistics of non-dummy variables. Table also contains summary statistics of the absolute ask rents and acquisition prices which are used as dependent variables. Sample mean, standard deviation, minimum and maximum are reported separately for distinct apartment subgroups: one-, two- and three-room apartments. The apartment type is listed in the second column of the Table 3.1.. Table 3.2 showcases the frequency tabulations of dummy variables by apartment types.

**Table 3.1 Descriptive statistics of non-dummy variables**

| Variable | Apartment type | Mean | Standard deviation | Minimum | Maximum |
|---|---|---|---|---|---|
| *Ask rent* | 1 | 718.31 | 100.91 | 400 | 1,250 |
| | 2 | 882.20 | 146.97 | 500 | 1,990 |
| | 3 | 1,112.73 | 292.74 | 550 | 3,990 |
| *Acquisition price* | 1 | 153,191 | 45,239 | 65,280 | 443,700 |
| | 2 | 192,480 | 66,522 | 80,634 | 559,980 |
| | 3 | 239,568 | 114,464 | 94,336 | 1,122,000 |
| *Apartment size* | 1 | 30.52 | 6.05 | 15 | 65 |
| | 2 | 50.65 | 7.47 | 27 | 108 |
| | 3 | 70.96 | 9.02 | 52 | 128 |
| *University* | 1 | 3.99 | 3.25 | 0.13 | 14.85 |
| | 2 | 4.89 | 3.25 | 0.10 | 14.85 |
| | 3 | 5.17 | 3.29 | 0.07 | 14.85 |
| *Metro* | 1 | 4.06 | 3.75 | 0.03 | 15.81 |
| | 2 | 4.98 | 3.74 | 0.02 | 16.34 |
| | 3 | 5.41 | 3.75 | 0.05 | 13.95 |
| *Railway* | 1 | 2.21 | 1.92 | 0.05 | 9.60 |
| | 2 | 2.74 | 2.32 | 0.06 | 9.76 |
| | 3 | 2.65 | 2.26 | 0.06 | 9.70 |

**Table 3.2 Frequency table of dummy variables**

| Variable | One-room apartment | Two-room apartment | Three-room apartment |
|---|---|---|---|
| *New apartment* | 47 | 161 | 73 |
| *Built 1962-1983* | 857 | 1,716 | 558 |

| | | | |
|---|---|---|---|
| *Built 1984-2016* | 354 | 1,114 | 332 |
| *Highest floor* | 498 | 902 | 275 |
| *Bottom floor* | 468 | 662 | 210 |
| *Sauna* | 96 | 824 | 307 |
| *Balcony* | 895 | 3,042 | 1,013 |
| *Owned lot* | 2,005 | 3,130 | 828 |
| *Elevator* | 1,250 | 2,083 | 521 |
| Number of obs. | 2,395 | 3,927 | 1,129 |

Moreover, as it is intuitively clear, most of the apartments in my sample are not either in the highest or lowest floor. Majority of the two- and three-room apartments have balconies, whereas it is rarer in one-room apartments. This can be explained by the increased space requirements for larger number of dwellers in two- and three-room apartments. Similarly, the presence of saunas is almost non-existent in one-room apartments, whereas large proportion of two- and three-room apartments have saunas. Across all subgroups, approximately half of the sample apartment buildings have elevators. Furthermore, majority of the observations across apartment types are on owned lots. This implies that only a minority of apartment blocks in HMA are built on land rented from e.g. the municipality.

As a conclusion, the sample is free of large outliers and inconsistent observations. Furthermore, the dataset has a good representation of all apartment types. Thus, I can conclude that the collected and matched sample is fit to analyse the disaggregated rental markets in HMA.

### 6.2 Evidence of apartment submarkets in HMA

The structure of this section is the following. Firstly, I present the double-log regression results for the whole aggregated sample. Secondly, I present the disaggregated double-log regressions for the hypothesized apartment submarkets. The analysis of coefficients from these regressions is conducted jointly after presenting my results in Table's 4 and 5.

Table 4 presents the double-log regression on the whole sample. Almost all of variables are significant in the one percent level in the model. Table 5 showcases double-log regressions for the distinct apartment submarkets. Similarly, in the submarket regressions, majority of the independent variables are significant. In the following, I discuss the results in detail and link them to earlier research results regarding rental markets.

# Table 4 – Double-log regression for aggregated sample

This table showcases regression results for the whole sample, consisting of all one-, two- and three-room apartments, employing a double-log regression model. The dependent variable is net rental yield in percentages. All independent variables are listed in the leftmost column. The rows with a count (# of dummy) are dummy variables. T-statistics are in parentheses; *** implies significance at the 0.01 level, ** implies significance at the 0.05 level and * implies significance at the 0.1 level.

| Variable | # of dummy | Coefficient (t-stat in parentheses below the coefficient) |
|---|---|---|
| *Apartment size* | | -0.060 (-10.22***) |
| *New apartment* | 281 | -0.083 (-8.48***) |
| *Built 1962-1983* | 3,131 | 0.087 (16.05***) |
| *Built 1984-2016* | 1,800 | -0.113 (-15.47***) |
| *Highest floor* | 1,675 | 0.006 (1.48) |
| *Bottom floor* | 1,340 | 0.017 (3.59***) |
| *Sauna* | 1,227 | -0.024 (-3.82***) |
| *Balcony* | 4,950 | -0.028 (-6.01***) |
| *Owned lot* | 5,963 | -0.050 (-9.66***) |
| *Elevator* | 3,854 | -0.008 (-1.93*) |
| *Constuction supply* | | -0.001 (-1.25) |
| *University* | | 0.026 (6.11***) |
| *Metro* | | 0.018 (4.35***) |
| *Railway* | | -0.007 (-2.09**) |
| *Rent year* | | Control variable |
| *Postal code area* | | Control variable |
| Constant | | 2.362 (25.87***) |
| Adjusted R-Squared | | 0.5790 |
| n | | 7,451 |

### Table 5 – Double-log regressions for apartment subgroups

The table showcases the double-log regressions for the apartment subgroups; one-, two- and three-room apartments. The dependent variables is net rental yield in percentages. There are 2,395 one-room apartments, 3,927 two-room apartments and 1,129 three-room apartments in my sample. Column on the left shows the independent variables, followed by columns of one-, two- and three-room apartments. Variables Rent year and Postal code area are controlled for all subgroups. Coefficients are presented without parentheses. T-statistics are in parentheses; *** implies significance at the 0.01 level, ** implies significance at the 0.05 level and * implies significance at the 0.1 level.

| Variable | One-room apartments | Two-room apartments | Three-room apartments |
|---|---|---|---|
| *Apartment size* | -0.240 | -0.097 | 0.068 |
| | (-16.54***) | (-5.90***) | (1.37) |
| *New apartment* | -0.081 | -0.095 | -0.057 |
| | (-3.95***) | (-7.65***) | (-2.26**) |
| *Built 1962-1983* | 0.064 | 0.106 | 0.117 |
| | (8.08***) | (13.93***) | (6.13***) |
| *Built 1984-2016* | -0.097 | -0.104 | -0.064 |
| | (-8.06***) | (-10.72***) | (-2.63**) |
| *Highest floor* | 0.014 | 0.007 | -0.008 |
| | (2.05**) | (1.28) | (-0.60) |
| *Bottom floor* | 0.012 | 0.016 | 0.020 |
| | (1.67*) | (2.40***) | (1.38) |
| *Sauna* | 0.036 | -0.026 | -0.056 |
| | (2.37***) | (-3.36***) | (-3.30***) |
| *Balcony* | -0.020 | -0.042 | -0.016 |
| | (-3.00***) | (-6.41***) | (-0.86) |
| *Owned lot* | -0.030 | -0.057 | -0.078 |
| | (-3.68***) | (-8.17***) | (-4.98***) |
| *Elevator* | 0.007 | -0.009 | -0.035 |
| | 1.27 | (-1.75*) | (-2.70***) |
| *Constuction supply* | -0.001 | -0.001 | 0.001 |
| | -0.56 | (-1.81*) | (0.46) |
| *University* | 0.021 | 0.031 | 0.017 |
| | (3.08***) | (5.38***) | (0.96) |
| *Metro* | 0.014 | 0.026 | 0.011 |
| | (2.32**) | (4.70***) | (0.75) |
| *Railway* | 0.002 | -0.002 | -0.020 |
| | 0.34 | -0.49 | (-2.15**) |
| *Rent year* | Control variable | Control variable | Control variable |
| *Postal code area* | Control variable | Control variable | Control variable |
| Constant | 2.636 | 2.247 | 1.731 |
| | (27.88***) | (31.44***) | (7.79***) |
| Adjusted R-Squared | 0.5797 | 0.6180 | 0.5780 |
| n | 2,395 | 3,927 | 1,129 |

*i. Apartment size*

Variable *Apartment size* is highly significant and has negative impact on net rental yield in the aggregated sample. This is along the lines of previous studies on the Finnish rental market, e.g. Moilanen and Terho (2010) who discovered that larger apartment's exhibit lower net rental yield. However, when examining the results for the submarkets, the results are different. The apartment's size is significant in one- and two-room apartments, but in three-room apartments, the sign is positive yet insignificant. This finding contradicts majority of papers on rental yield determinants, which often raise size as the most important variable (See e.g. Malpezzi, 2003; Kain and Quigleu 1970).

However, majority of the past studies have focused simply on the relationship between rent and price. For example, Garner and Verbrugge (2009) and Tian (2008) discovered that rents are concave with apartment prices, i.e. rents increase at a decreasing rate relative to value. My findings indicate that this relationship does not hold in three-room apartments. Furthermore, increasing apartment size has high negative impact on net rental yield in the one-room and two-room subgroups, implying that increasing apartment size has negative impact on net rental yield only in the more compact one- and two-room apartments. For example, increasing the size of one-room apartments with one percent will lead to 0.24 percent decrease in the net rental yield. This is rather substantial difference as the similar differential in two-room apartments is only 0.09 percent. One explanation for the phenomenon is the price of the smaller apartments. In the Helsinki area, smaller apartments often have the highest price per square meter driving the return on these spaces low. Second feasible explanation is that rental dwellers in smaller apartments are not ready to pay extra premium on the rent for the extra space after the apartment has all the required amenities. For example, two- and one-room apartments can have the same amenities, but the apartment size can vary significantly due to poor planning. The rental dwellers are probably not willing to pay similar rents per square meter for the extra space, when compared to the mandatory space with required amenities.

*ii.* *New apartment*

Variable *New apartment* has a negative and significant, at the one percent level, impact on net rental yield in both the aggregated sample and all subgroups except three-room apartments, where the significance is present at 5 percent level. Hence, new apartments command lower net rental yields when compared to older apartments in HMA. In general, new apartments are often sold with high prices due to high construction costs in HMA. Rental dwellers seem to be reluctant to pay higher rents in relation to the higher acquisition price, translating into lower yield. However, acquiring new apartments as investments is relatively popular for rental investors due to the specific transaction structure available for new constructions in Finland. Purchasers of new apartments can often pay only certain percentage of the apartment's value in cash and remainder of the transaction value is paid in the form of housing company loan payments over time – these loans usually have lower interest rates than what the individual investors are able to gain from their own banks. Furthermore, housing company loan payments are tax deductible in most cases[7] and often the first few years are installment free, further increasing the appeal of these transactions. When it comes to my results, my findings contradict the findings of Melakari (2014), who found that newer apartments command higher net rental yields when compared to older apartments in HMA. This can be explained by my research setting of using the full transaction price, not only the equity portion or the ask price, as the basis for net rental yield calculations, which reduces the possible bias from characteristics of new apartments.

*iii.* *Built 1962-1983*

Apartments built between years 1962 and 1983, in the era when urbanization trend was beginning in Finland (see Section 2.3 for more information), which are often seen as low quality buildings for owner-occupancy, have positive and statistically significant impact, at one percent level, on the net rental yield for all apartment subgroups. One possible explanation is that there is a divergence of preferences between the owner-occupants and rental dwellers. As owner-occupants are not prepared to pay high prices for these apartments, rental dwellers could be less stringent on the quality of apartments they seek for short-term dwelling, thus these apartments have positive impact on the net rental yield in general. Laakso's (1997) findings back this argument, as he found that apartment buildings built between 1960's and 1970's are valued lower when compared to other apartments, *ceteris paribus.*

*iv.* *Built 1984-2016*

Apartments built after the 1984's are characterized as having a negative impact on the net rental yield of these apartments. All subgroups except three-room apartments have negative and statistically significant coefficient at the one percent level. The feasible explanations for this phenomenon is the opposite of the hypothesis explaining why apartments built in the 1962-1983 have positive impact on the obtainable net rental

[7] These payments are tax deductible if they are recognized in the P&L of the housing company instead of capitalized on the balance sheet

yield. Firstly, the apartments built after 1983 might be trading on a relatively higher price when compared to other apartments, as they were built to provide more utilities than mere accommodation for dwellers. Second explanation is linked to the economics of complete acquisition price: newer apartments are often free of large renovations in the near-term future. Even by discounting the cost of these renovations to the purchase price would have only small impact on the purchase price, when compared to older apartments that are facing costly pipe or façade repairs in the near-term future. In other words, my dataset most likely has more truthful image of the purchase price of apartments built between 1984 and 2016, when compared to apartment built between 1962 and 1983, due to the lack of data regarding upcoming renovations and the resulted inability to compute the impact they have on net rental yield.

*v.* *Highest floor*

This variable is insignificant in the aggregated sample, indicating that apartments in the highest floor offer no impact on the rental yield profile of the apartment. However, when looking at the subgroups, highest floor variable is significant and positive at the 5 percent level in the subgroup of one-room apartments. It is insignificant in other subgroups. This implies that in general, one-room apartments in the highest floors are more appealing to rental dwellers and they are willing to pay relatively higher rents for these apartments.

*vi.* *Bottom floor*

When looking at the aggregated sample, apartments in the bottom floors of the buildings have significantly, at one percent level, higher net rental yields than other apartments. However, the story is different in the subgroups. Only the one– and two– room subgroups have similar effect, whereas *Bottom floor* has no significant impact on net rental yield in subgroup of three-room apartments. One explanation is that larger apartments, i.e. three-room apartments, are not as appealing on the lower floors for the rental dwellers. One explaining hypothesis is that as the family-size often increases in three-room apartments, these tenants value the increased privacy of the apartments in higher floors to larger extent than the tenants in one- and two-room apartments.

*vii.* *Sauna*

In the full double-log regression model, having a sauna in an apartment has a negative and highly significant, one percent level, impact on the net rental yield. Interestingly, when looking at the subgroups, similar results are showcased only in the subgroup of two- and three-room apartments. In the subgroup of one-room apartments, the coefficient is positive and significant at the one percent level. This can be explained by two factors, often the price of sauna's is taken into account in the acquisition price of apartments, but most likely not all rental dwellers value this utility in their apartments, thus decreasing the net rental yield. Secondly, the fact that my sample has only 96 apartments with saunas in one-room apartments when compared to the number of one-room apartments in my sample at 2,395 might distort these results. The relative number of saunas in other apartments types is substantially higher.

*viii.* *Balcony*

Having a balcony in the apartment has a negative and highly significant (at one percent level) impact on the net rental yield in the aggregated sample. However, the impact is two-fold in the subgroups. In both one- and two-room apartments, the impact of balcony is negative and significant at one percent level on the net rental yield. However, the results are different in three-room apartments. In this subgroup, the variable is non-significant.

*ix.* *Owned lot*

The variable owned lot is negative and highly significant, at the one percent level, on both the aggregate sample and the distinct subgroups. The pro-rata distributed price of the lot in apartment prices most likely drives down net rental yield in these observations. In the case of rental lots, the cost of the lot should be included in the maintenance charges. One could assume that the cost should be approximately equal in buy or rent situation scenarios. Given how significant and negative the impact of owning the lot is, it leads us to a hypothesis that the rental levels on rented lots are much lower than what the market prices of these lots would implicate. One plausible explanation for this phenomenon is that often, at least on lots owned by municipalities, the lot rental contracts can be fixed for decades, often for 50 years. If these contracts are closing maturity and the pricing is not indexed, the rental level of the lots is probably significantly lower from current market levels due to general inflation.

*x.* *Elevator*

The results for elevator are rather surprising. In the aggregate sample, the subgroup for two-room and the subgroup for three-room apartments the variable is significant and negative. One possible hypothesis for these results is that the difference is caused by maintenance charges from operating the elevator. In Finland, these are charged to dwellers in the form of maintenance charges based on the size of the apartment. For this reason, the maintenance fee can be quite substantial for two- and three-room apartments when compared to small one-room apartments, causing the coefficient elevator to be negative and significant in the former subgroups.

xi. *Construction supply*

Against my hypothesis two, construction supply is insignificant in explaining net rental yield in the aggregate sample. However, it is significant at the 10 percent level in the subgroup of two-room apartments. Without having more granular information about the type of apartments developed in these areas the plausible explanation is that most of the developed apartments in HMA are two-room apartments – which has a negative impact on the chargeable net rental yield from these apartments due to increased supply.

xii. *University*

The distance to university is positive and significant, at one percent level, in all regressions except the subgroup for three-room apartments. As the variable reflects the distance to closest university campus in kilometers, this finding implies that as the distance to university campus increases, the rental yield increases. This finding is aligned with the findings of Melakari (2014), who discovered that rent levels are lower closer to university campuses.

xiii. *Metro*

Close proximity to metro is also highly significant and positive, at one percent level in one-room apartments and aggregated sample and at 5 percent level in two-room apartments, in all other groups except three-room apartments where the variable is insignificant. When it comes to three-room apartments, the finding can be explained by the fact that often the tenants in three-room apartments are often families with children and thus rental decision is most likely weighted on other criteria, e.g. close proximity of schools and safe outdoor areas. In regards to the positive and highly significant variable for one – and two-room apartments, findings are aligned with Laakso's (1997) findings, which indicated that the positive impact of accessibility does not outweigh the possible negative externalities from the vicinity metro stations.

xiv. *Railway*

To the contrary from the metro station coefficients, the coefficient of railway is significant and negative in the 5 percent level in the aggregate sample. However, this effect seems to be driven by the three-room sample where the coefficient is negative and significant at the one percent level. In other samples, the coefficient is insignificant. This implies that for dwellers in three-room apartments, the positive impact of accessibility outweighs the negative externalities from the close vicinity of railway stations, contradicting Laakso's (1997) findings.

The full regression model has 7,451 observations, whereas the segmented regressions for one-, two- and three-room apartments have 2,395, 3,927 and 1,129 observations respectively. The adjusted R-squared is 0.579 in the aggregated regressions and 0.579, 0.618 and 0.578 in the respectable one-, two- and three-room regressions. The adjusted R-squared coefficients imply that the variables explain the variation in the model to the best degree in the sample for two-room apartments. However, the adjusted r-squared is on a relatively high level in all regressions, implying high and consistent model fit for all subgroups. Furthermore, the constant is highly significant at the one percent level and positive in all regressions. When taking into account the economic magnitude of the variables, most of the coefficients have only marginal effect on the net rental yield. Most of the independent variables increase or decrease the net rental yield in small percentages, which is explained by the by default small net rental yield numbers. In addition, to ensure that the validity of the results is not reliant on the methodology, and for easier interpretation, I also ran distinct regressions for aggregate sample and subgroups with linear-linear specification. These results can be found from the appendix in Table 10. The results are similar as in the double-log specified model.

In general, my empirical results clearly indicate that the apartment rental markets are disaggregated in HMA. Assessing rental apartments as an aggregate mass might lead to un-optimal investment decisions, as different apartment characteristics might have positive or negative impact on the net rental yield based on the type of apartment.

However, to validate these results, it is important to examine whether the slopes of the regressions and the variables coefficients in the one- two- and three-room subgroups are statistically significantly different from each other to confirm the disaggregation of apartment rental markets in HMA. Hence, Section 6.3 is focused on examining which of the apartment subgroups differ significantly from each other and specifically, which coefficients statistically differ between the apartment subgroups.

## 6.2 Significance of model differences

In this section, I further evaluate the existence of disaggregated apartment rental markets in the HMA. The previous section showcased how the coefficients and their significances varied between the apartment subgroups. Hence, backing my hypothesis of disaggregated rental markets in HMA from the viewpoint of economic significance. This part of the analysis starts with a Chow-test, showcasing if the regressions slopes in the subgroups of one-, two-, and three-room apartments statistically significantly differ from each another. This test verifies the observed difference between the sub-sample regressions statistically. Secondly, I will examine which of the coefficients in the subgroups are causing the differences between the subgroup models with Tiao-Goldberger F-test. Finally, I will discuss how my empirical results reflect on my initial hypotheses.

Firstly, to formally test the null hypothesis that the slopes of the subgroups regressions are similar, i.e. the apartment rental market is not disaggregated, a series of Chow (1960) tests are conducted on the apartment subgroups. First step of the analysis is to estimate the aggregate models for a combination of one-room and two-room apartments, a combination of one-room and three-room apartments and two-room and three-room apartments. For the sake of clarity, I omitted reporting the results of this step. After this, I derived the Chow-test statistics to test the null hypothesis that the rental markets in HMA are not disaggregated. The formula to calculate the statistic is explained in detail in Section 4.2.1. The results can be found from the Table 6**.** These results indicate that the apartment rental market is disaggregated in the HMA and I can reject the null hypothesis of the Chow-test, that the regressions models would be statistically similar.

**Table 6 - Chow-test**

This table showcases the Chow-test results. The null hypothesis is that the slopes of one-room, two-room and three-room apartment models are similar. The statistic is retrieved by running distinct aggregated regression models for all available pairs. F-statistics are in parentheses; *** implies significance at the 0.01 level, ** implies significance at the 0.05 level and * implies significance at the 0.1 level.

| Regression pairs | Chow-test statistic |
|---|---|
| One-room, Two-room | 7.76*** |
| Two-room, Three-room | 6.28*** |
| Two-room, Three-room | 3.07*** |

Secondly, it is important to assess which of the coefficients are statistically significant from each other between the apartment subgroups. To examine this, I employ Tiao-Goldberger F-test (please see Section 4.2.2. for more information) to test the statistical difference between sub-sample regression coefficients.

Table 7 summarizes the findings from the Tiao-Goldberger test. Variables *Apartment size*, *Built 1962-1983*, *Built 1984-2016*, *Sauna*, *University* and *Metro* are statistically significant at the one percent level between the three subgroups. Variables *Owned lot* and *Construction supply* are statistically significant at the five percent level between the three subgroups. In other words, these 8 variables are driving the observed difference between net rental yield functions in one-, two- and three-room apartments, whereas the impact of variables *New apartment, Highest floor, Bottom floor, Balcony and Railway* on net rental yield is statistically insignificant between the samples.

**Table 7 - Tiao-Goldberger test**

This table showcases the Tiao-Goldberger F-test results. The null hypothesis of Tiao-Goldberger test that the coefficients are same. The leftmost column reports the independent variables. The columns labeled as 'Coefficients' include the coefficients of the variables from the original double-log regressions from Table 4 and 5. Statistical significances of these variables are not report. The final column on right hand side showcases the Tiao-Goldberger F-test results; *** implies significance at the 0.01 level, ** implies significance at the 0.05 level and * implies significance at the 0.1 level.

| | Coefficients | | | |
|---|---|---|---|---|
| **Variable** | **One-room apartments** | **Two-room apartments** | **Three-room apartments** | **Tiao-Golberger F-test statistics** |
| *Apartment size* | -0.240 | -0.097 | 0.068 | 41.63*** |
| *New apartment* | -0.081 | -0.095 | -0.057 | 2.18 |
| *Built 1962-1983* | 0.064 | 0.106 | 0.117 | 25.10*** |
| *Built 1984-2016* | -0.097 | -0.104 | -0.064 | 25.53*** |
| *Highest floor* | 0.014 | 0.007 | -0.008 | 1.94 |
| *Bottom floor* | 0.012 | 0.016 | 0.020 | 0.96 |
| *Sauna* | 0.036 | -0.026 | -0.056 | 20.79*** |
| *Balcony* | -0.020 | -0.042 | -0.016 | 1.52 |
| *Owned lot* | -0.030 | -0.057 | -0.078 | 4.40** |
| *Elevator* | 0.007 | -0.009 | -0.035 | 10.74 |
| *Constuction supply* | -0.001 | -0.001 | 0.001 | 4.50** |
| *University* | 0.021 | 0.031 | 0.017 | 6.83*** |
| *Metro* | 0.014 | 0.026 | 0.011 | 5.05*** |
| *Railway* | 0.002 | -0.002 | -0.020 | 2.30 |
| *Rent year* | Control variable | Control variable | Control variable | |
| *Postal code area* | Control variable | Control variable | Control variable | |
| Constant | 2.636 | 2.247 | 1.731 | |
| Adjusted R-Squared | 0.5797 | 0.6180 | 0.5780 | |
| N | 2,395 | 3,927 | 1,129 | |

Finally, to assess my initial research hypotheses, Tiao-Goldberger test results verify the prior findings from the comparison of separate subgroups regressions and Chow-test, both indicating that the apartment rental market is disaggregated in HMA between one-, two- and three-room apartments. With these empirical results, I conclude that I fail to reject hypothesis one. The rental yield functions in HMA are statistically significantly different. Moreover, the important coefficients determining the rental yield between one-, two- and three-room apartments vary between the subgroups.

However, I partially reject hypothesis two, stating that construction supply in postal code area has adverse impact on net rental yield in apartments in same postal code area. The variable has only significant effect on two-room apartments. There are two different explanations for this result. Firstly, my dataset does not contain information regarding the type of apartment's constructed, i.e. I cannot compare the net rental yields for one-room apartments directly to supply of one-room apartments. Secondly, the construction projects are not directly tied to postal code areas. It is likely that the impact these developments have is spread also to neighboring postal code areas.

Finally, at this point I also fail to reject my initial hypothesis number three, stating that the previously documented inverse relationship between net rental yield and apartments size is driven by the subgroups of one- and two-room apartments. My findings indicated that the negative relationship is visible only in one-room and three-room apartments. Moreover, the difference between coefficients is statistically significant between the subgroups. However, to further assess the relationship between apartment types and the size of apartments, I conducted extra analysis to verify these results. This analysis will be presented in Section 6.5.

However, before assessing hypothesis three in detail, I expand my analysis of disaggregated apartment markets from net rental yield to absolute net rents and apartment prices. In essence, net rental yield is formed by the relationship between apartment prices and absolute net rents. Hence, it is important to first explore whether the model coefficients influence apartment prices and absolute net rents with different magnitudes. Such elasticities would clarify whether the price differential associated with certain apartment characteristic, such as larger apartment size, is translated into equal impact on absolute rents.

### 6.4 Relationship between apartment prices and rents

In addition to examining the determinants of net rental yields in HMA, I employed similar double-log regressions on the apartments prices and absolute net rents from the apartments to examine the relationship between apartment prices and absolute rents further. These regressions can be found from the appendix in Tables 11 and 12. Firstly, I will briefly discuss the results of apartment price regressions for aggregated and disaggregated samples and reflect these findings to my previous results. Secondly, I will discuss the relationship between price, absolute net rents and net rental yield to further assess whether the elasticity of prices or rents influences net rental yield formation in the subgroups of one-, two- and three-room apartments.

When looking at the aggregate model, the findings in general are rather intuitive and inverse of the net rental yield coefficients. First of all, the coefficients for new apartments size and new apartments are highly significant, at one percent level, and negative. These higher prices can partly explain the reduced net rental yield from these apartments. Furthermore, the apartments built between 1962 and 1983 are sold on a discount when compared to other apartments. The coefficients are significant at the one percent level, which is also inversely related to net rental yield coefficient for this variable. Apartments sold in the bottom floor are sold at a discount and the coefficient is highly significant at the one percent level. However, the net rental yield in these apartments is also higher, so the price does not seem to be the explaining factor. Coefficients *Sauna, Balcony, Owned lot* and *Elevator* are all highly significant at the one percent level and positive. However, all coefficients are negative in the net rental yield regressions. This can be explained by the different preferences of owner-occupants and rental dwellers. *Construction supply* is highly significant and positive in the aggregate price regression. However, this variable is not significant in the aggregated net rental yield model. As for the location related variables, the dummy variables for *University* and *Metro* are negative and significant at the one percent level in the aggregated regression. These coefficients are positive and highly significant in the price regressions. This implies that the close proximity of universities and metro stations has positive impact on apartment prices, but the rental dwellers do not seem to value these characteristics to compensate the higher acquisition costs.

Moving to disaggregated price regressions, variable *Apartment size* is highly significant, at one percent level, and positive in all price regressions. This is intuitive as the total price of the apartment increases with the total space. The magnitude of this effect diminishes moving from one – to three-room apartments, showcasing how the average price per square meter decreases as the apartments get bigger. The dummy variables for *New apartment*, *Built 1962-1983* and *Built 1984-2016* are, similarly to the aggregate model, highly significant at the one percent level and positive. This mirrors the coefficients in the disaggregated net rental yield models. In general, the rest of the findings in the disaggregated model follow this trend and largely mirror the significances in the disaggregated net rental yield models. This observation fortifies the hypothesis established in previous section that owner-occupants and rental dwellers have highly differing preferences when it comes to their dwelling characteristics.

To conclude, it seems that the higher price of certain apartments can explain the difference in net rental yield. This applies for example to new apartments and apartment built between the years 1962 and 1983. Apartments with certain characteristics are clearly more preferred by owner-occupants when compared to rental dwellers. For example, owner occupants seem to prefer apartments further away from University campuses whereas rental dwellers clearly prefer these apartments and hence these apartments command higher net rental yield percentages *ceteris paribus*

However, as previously discussed, one possible explanation for the inverse relationship between apartment prices and net rental yields is that the model coefficients impact absolute rents and apartment prices with differing magnitude. In the following, I will review whether this elasticity effect is driving the observed inverse relationship between apartment prices and net rental yields.

Table 8 below summarizes the impact the independent variables have on apartment prices, absolute net rents and net rental yield. The positive and negative signs in the table imply that these variables have positive or negative impact on the respectable dependent variable. Elasticity between price and absolute rents can be observed if independent variables have similar impact on absolute rents and prices but the impact on net rental yield is inverse. As an example, if variable *Apartment size* increases both prices and absolute rents but the impact on net rental yield is negative it is due to elasticity effect. In other words, increasing size increases the apartment price in larger magnitude than the rent, resulting in reduced yield for the investor.

**Table 8 - Relationship between rental yield, apartment prices and absolute rents**

This table showcases the key takeaways of the regression analyses presented in Tables 5, 11 and 12. First row presents the subgroups, second row presents the dependent variable in the regression and first column represents the independent variables. Sign + implies that the dependent variable has statistically significant effect on the dependent variable. Sign – implies that the independent variable has statistically significant and negative effect on the dependent variable. Sign 0 implies that the independent variable does not have statistically significant impact on the dependent variable. Elasticity effect implies that the variables have the expected sign, but the magnitude is larger for prices. This is the case for numerous variables and is explained by the discovery that the prices in my sample seem to be more elastic to both negative and positive attributes than the apartment rents across all subgroups.

| | One-room apartments | | | Two-room apartments | | | Three-room apartments | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Variable** | **Yield** | **Rent** | **Price** | **Yield** | **Rent** | **Price** | **Yield** | **Rent** | **Price** | **Key takeaway** |
| *Apartment size* | - | + | + | - | + | + | 0 | + | + | Impact of elasticity effect |
| *New apartment* | - | + | + | - | + | + | - | + | + | Impact of elasticity effect |
| *Built 1962-1983* | + | - | - | + | - | - | + | - | - | Impact of elasticity effect |
| *Built 1984-2016* | - | + | + | - | + | + | - | + | + | Impact of elasticity effect |
| *Highest floor* | + | 0 | - | 0 | 0 | 0 | 0 | 0 | + | Differing preferences between tenants and owner-occupants |
| *Bottom floor* | + | 0 | - | + | - | - | 0 | 0 | 0 | Impact of elasticity effect |
| *Sauna* | + | 0 | - | - | + | + | - | 0 | + | Impact of elasticity effect or differing preferences |
| *Balcony* | - | + | + | - | - | + | 0 | - | 0 | Impact of elasticity for one-room apartments |
| *Owned lot* | - | + | + | - | + | + | - | 0 | + | Impact of elasticity effect |
| *Elevator* | 0 | + | 0 | - | + | + | - | + | + | Impact of elasticity effect |
| *Constuction supply* | 0 | 0 | 0 | - | + | + | 0 | 0 | 0 | Impact of elasticity effect in two-room apartments |
| *University* | + | - | - | + | - | - | 0 | 0 | 0 | Impact of elasticity effect for one-, and two-room apartments |
| *Metro* | + | - | - | + | - | - | 0 | 0 | 0 | Impact of elasticity effect for one-, and wo -room apartments |
| *Railway* | 0 | - | - | 0 | - | 0 | - | 0 | + | Mixed results; not valued by three-room rental tenants |

To conclude, large majority of the observed sign differences between net rental yield and apartment price coefficients seem to be driven by the elasticity effect, i.e. that the apartment prices increase or decrease relatively more when compared to the respectable increase or decrease in absolute rent levels.

However, it is important to keep in mind that apartment markets are affected by multiple factors beyond the apartment characteristics, which might have an impact on the observed elasticities between prices and rents. For example, Einiö et al. (2008) discussed the behavioral factors affecting the selling prices of apartments in HMA. They concluded that in HMA apartment sales market, sellers are often loss-averse and seek to sell the apartment at least with the value it has been purchased with. Similar behavioral triggers might be present more universally in the apartment sales and rental markets, causing the observed elasticities between apartment price and absolute rent levels. Landlords might exhibit from similar behavioral biases. For example, they might be setting the apartment rents close to their individual break-even level, where rent payments cover apartments operating expenses, debt service and capital income taxes, instead of charging market rents. Hence, the joint effect of these behavioral factors might distort the relationship between apartment prices and rents.

### 6.4.1 Relationship between apartments size and type

As discussed in the previous section, my findings regarding the importance of apartment's size contradicted the overall academic consensus that size is one of the most important determinants to rental yield (see e.g. Malpezzi, 2003 and Kain and Quigleu 1970). My results indicated that the impact of apartment's size is not universal, and it is in fact insignificant in the three-room subgroup, aligned with my hypothesis 3. However, to further assess the impact variable *Apartment size* has on the rental yield, and to further examine my hypothesis 3, I conducted an extra analysis to assess the impact apartment's size has on net rental yield. This is conducted with a regression model, where the apartment types are controlled with dummy variables and introducing interaction variables between apartment's size and apartment type in to the model. Results of this test are visible in Table 9 below.

**Table 9 - Relationship between apartments size and type**

This table showcases the double-log regression on net rental yield with following independent variables: size of the apartment, dummy variables for one-, two- and three-room apartments and interaction terms between apartment size and the aforementioned dummy variables. The dependent variable is the net rental yield in percentages. The purpose of this regression is to analyse the distinct impact of size and size in relation to apartment's type to further assess how the size impacts the net rental yield between different apartment types. Coefficients are presented without parentheses. T-statistics are in parentheses; *** implies significance at the 0.01 level, ** implies significance at the 0.05 level and * implies significance at the 0.1 level.

| **Variable** | **Coefficient (t-stat in parentheses below the coefficient)** |
|---|---|
| *Apartment size* | -0.182<br>(-7.81***) |
| *One-room apartment* | 0.641<br>(2.62***) |
| *Two-room apartment* | -0.461<br>(-1.85*) |
| *One-room size* | *Omitted* |
| *Two-room size* | 0.302<br>(9.09***) |
| *Three-room size* | 0.192<br>(3.24***) |
| Constant | 1.736<br>(7.49***) |
| Adjusted R-Squared | 0.0132 |
| n | 7,451 |

Table 9 offers few key insights into the relationship between apartment size and apartment's type. I will first describe the key observations and then assess the robustness of the results and how they reflect on my hypothesis 3.

Firstly, increasing size seems to universally decrease rental yields. Secondly, control variables for apartment types, one-room apartment and two-room apartment, have both economically and statistically highly significant impact on the net rental yield. However, the sign is positive for *One-room apartment* and negative for *Two-room apartment*. Furthermore, the interaction variables *Two-room size* and *Three-room size* are both statistically significant and positive at one percent level. Furthermore, the economic impact of *Two-room size* interaction variable is larger than the impact of *Three-room size* interaction variable.

These observations have interesting implications. Firstly, they further solidify the hypothesis of disaggregated apartment rental markets in terms of apartment size. There are clearly statistical differences between the importance of size and amount of room's apartment has. Secondly, apartment size in relation to apartment type has larger impact on net rental yield on two-room apartments than three-room apartments. This could imply that apartment modification transforming smaller three-room apartments into larger two-room apartments could be a profitable investment opportunity, given the other apartment characteristics are favorable.

However, the overall fit of the model is rather low with adjusted R-squared of 0.0132 and as the model omits variable *One-room size* due to multicollinearity with other variables. Despite these factors, the extra analysis of the importance of apartment's size showcases how the impact of apartment's size does in fact differ between apartment types.

Hence, I conclude that that I fail to reject my initial research hypothesis 3, the concave relationship between apartment size and rental yield is largely driven by subgroups of one- and two-room apartments.

### 6.5 Extra considerations and robustness analysis

In this section, I will discuss the robustness of my results. Firstly, I discuss the impact of postal code area of the apartment has on the net rental yield. Secondly, I discuss the impact socio-economic factors have on the net rental yield, i.e. the unemployment rate and median income. Thirdly, I discuss possible multicollinearity between the independent variables in my regression models.

#### 6.5.1 Impact of postal code areas on net rental yield

In the previous regressions, I have controlled the importance of location and focused on other tangible independent variables. In this section, I discuss the significant control variables for the postal code areas and their impact on net rental yield. Table 13 in the appendix highlights the coefficients and t-statistics for the significant postal code area variables.

As discussed, the postal code area division is based on the third significant number of the postal codes, which are based on international standards. Postal codes starting as 01 are in Vantaa region, postal codes starting with 02 are in Espoo region. Rest are from Helsinki region. Interestingly, most of the significant postal code area variables are negative. The only region where location has significantly positive impact on net rental yield is 01200 area in Vantaa, Hakunila for two- and three-room apartment and 01400 area in Vantaa, Rekola for all apartment subgroups. This implies that there exists only two postal code areas from which rental investors can expect higher net rental yield, *ceteris paribus,* when compared to other postal code areas. In other areas, the impact is negative or insignificant and other apartment characteristics possess higher explanatory power in the formation of net rental yield.

### 6.5.2 Socio-economic factors on postal code levels

Multiple studies have reported the importance of socio-economic factors in defining the rental yields or apartment prices (see e.g. Juntto 2007; Fujita, 1989; Li and Brown, 1980). Due to lack publicly available data, I had postal code specific information regarding the unemployment rate and median income levels only between the years 2012 and 2014. This data was retrieved from the publicly available PAAVO-database, from Statistics Finland. To keep the main sample size as large as possible, socio-economic factors are not part of the main analysis, but their impact on net rental yield will be discussed in the following robustness analysis section.

The results containing the socio-economic variables can be found from the Table 13 from the appendix. The regression only contains observations from the time period 2012-2014. Contradicting the previous literature, both variables *Unemployment rate* and *Median income* are insignificant in almost all of the apartment type subgroups. The only subgroup where the variable median income was statistically significant at the 10 percent level is the subgroup of three-room apartments. There are two plausible explanations for these results, first on is the limited geographical scope in my study and the second is the limited time frame in my study.

Firstly, I am only studying the Helsinki metropolitan area, which decreases the variation in socio-economic statuses between the neighborhoods. The explanatory power on net rental yield would most likely be higher in a sample covering the full country. For example, Laakso (1997) discovered that socio-economic factors of the neighborhoods have high impact on the price levels of the apartments. My findings indicate that the rental yield stays stable across areas with different socio-economic standings within HMA.

Furthermore, with a sample spanning only a few years' time-period, the larger variation and developments in the socio-economic factors and their impact on net rental yield remain unverified. With these results in mind, I can conclude that the reported socio-economic factors do not have large impact determining the net rental yield of HMA rental apartments.

### 6.5.3 Multicollinearity and correlation between variables

As a robustness analysis, I also conducted a pairwise correlation analysis to discover highly correlated variables to identify possible multicollinearity in my regressions. The correlation matrix is visible in the appendix, in Table 15.

In general, the correlation between my main independent variables is small. Even the largest correlation coefficient is below 0.5 level. However, almost all correlation coefficients are statistically significant at either 5 percent or one percent levels. This is not surprising given the nature of the variables.

To give few examples about the intuitive nature of the higher correlation coefficients, there are two practical examples. Variables *Built 1962-1983* and *Built 1984-2016* exhibit high correlation with the variable *Sauna*. The nature of correlation is negative for *Built 1962-1983* and *Sauna* and positive for *Built 1984-2016* and *Sauna*. This implies that the trend of constructing apartment with saunas has shifted after the 1983. Second example is the negative and highly significant correlation between the distance to metro and railway stations. Glancing at the metro and railway maps of Helsinki metropolitan area reveals the reason for this phenomenon; there is only small overlap between the metro and railway network. Due to the intuitive nature of the correlation coefficients and lack of large magnitudes of correlation between my main independent variables, I do not proceed with extra analyses due to the results from the correlation matrix.

# 7. CONCLUSION

This study revealed that the rental markets in the Helsinki metropolitan area are heterogeneous. The aggregate model and distinct regressions for one-, two- and three-room apartments implied significant differences in the slopes of the subgroup regressions and in the net rental yield coefficients between apartment types.

Despite the financial benefits of real estate diversification (See e.g. Hartzell et al., 1986) the academic research has largely focused on aggregate rental markets (See e.g. Clapp et al., 1992). Only a handful of studies have explored the existence of rental submarkets or segmented rental markets according to apartment type. However, the studies dividing apartment markets into submarkets have found statistically significant differences based on either location (See e.g. William, 1996) or apartment types (See e.g. Wolverton et al. 1999). However, in Finnish context this is the first study discussing the potentially diverging rental yield determinants between apartment subgroups based on the amount of rooms in the apartment. Previous studies have focused on price determinants (See e.g. Laakso, 1997) or studying rental markets as aggregate sample (Terho and Moilanen, 2010).

The research question of this thesis was aimed to fill this void. The overall contribution to the existing literature is three-fold. Firstly, my thesis was the first paper, to my best knowledge, examining the disaggregation of apartment rental markets in Finland. The concept of disaggregated apartment rental markets has large-scale implications for the rental investors. As this study highlighted, different apartment types should be regarded as different type of investments due to the differing rental yield determinants between apartment types. Making investment decisions based on aggregated rental market model might lead un-optimal investment decisions. Secondly, this study uniquely combined actual transaction data to rental advertisements. Majority of previous studies have been conducted with datasets employing apartment ask prices instead of actual prices. As I had access to transaction data with realized sales prices, I was able to assess net rental yield determinants based on a non-biased sample in regards to differences between ask prices and realized prices. Thirdly, my thesis is the first study in Finnish context incorporating a supply side variable, specifically apartment supply in a postal code area, into the equation when assessing rental yield.

The empirical analysis in this study was conducted with a novel dataset. I received apartment rental advertisements from Oikotie, which is part of Sanoma Oyj. Data regarding apartment transactions with realized sales prices was received from KVKL. In addition to these data sources, various independent variables, e.g. apartment coordinates and postal code level apartment supply, were received from HSL, Statistics Finland and Google Maps. Datasets of actual transactions and rental advertisements were matched with exact addresses and apartment various apartment characteristics. Moreover, I manually examined the sample to ensure the data quality. The final sample consisted of 7,451 rental apartments, of which 2,395 were one-room apartments, 3,927 were two-room apartments and 1,129 were three-room apartments.

The methodology employed in my study follows the methodology from previous literature (See e.g. Wolverton et al. 1999). Firstly, I compared the regression coefficients from the aggregated and disaggregated sample and discussed the differences in rental yield determinants between the apartment subgroups. Secondly, I examined the statistically significant differences between these regressions with a Chow-test. Thirdly, to examine which coefficients are causing the observed differences between the subgroup regressions, I employed a Tiao-Goldberger F-test. Finally, I conducted multiple tests to ensure the robustness of my results by e.g. examining the relationship between rental yield, apartment prices and absolute rents.

My results clearly indicated that that the apartment rental markets in HMA are disaggregated. The coefficients determining net rental yield between aggregated and disaggregated samples had statistically significant magnitudes and signs. For example, variable *Apartment size* had statistically significant and negative impact on net rental yield in subgroups of one- and two-room apartments but the variable was insignificant in the subgroup of three-room apartments. Moreover, the results from Chow-test solidified these results; the regression slopes between the subgroups of one-, two- and three-bedroom apartments were statistically significantly different. Furthermore, Tiao-Goldberger test revealed that variables *Apartment size*, *Built 1962-1983*, *Built 1984-2016*, *Sauna*, *Owned lot*, *Construction supply*, *University* and *Metro* were causing the statistically significant difference between the subgroup regressions. Thus, these eight variables have significantly different impact on net rental yield between one-, two- and three-room apartments.

When taking these results to the context of my initial research setting and hypotheses, the results were ambiguous. Firstly, I failed to reject my hypothesis number one. According to my empirical results, the apartment rental market in HMA is clearly heterogeneous. Secondly, the results regarding my second hypothesis, regarding the impact apartment supply has on net rental yield, were mixed. The variable was insignificant explaining net rental yield in all subgroups except in the subgroup of two-room apartments. However, according to Tiao-Goldberger test, this variable was one of the root causes behind the differing rental yield equations between apartment subgroup. Finally, in the light of my third hypothesis, that larger apartment size has adverse effect on net rental yield only in one- and two-room apartments, I failed to reject my initial hypothesis. The variable *Apartment size* was insignificant in the three-room subgroup regression, but significant and negative in the subgroups of one- and two-room apartments.

However, due to the limited amount of academic papers regarding the rental markets in Finland, there are still multiple areas for further research. First venue is linked to the ambiguous results related to relationship between apartment size, apartment prices and rents. As discussed, the rental market price setting might experience from undocumented behavioral patterns, e.g. loss-aversion of rental investors, similar to owner-occupants selling their homes, which might drive investors to set their rental levels close to arbitrary breakeven points. For example, investors might have tendency to set the rents close to a level which offsets the mortgage payments, interest payments and all maintenance charges from the apartment instead of the charging market rents. Examining these undocumented behavioral biases related to rent setting has the potential to further explain the relationship between apartment prices, absolute rents and net rental yield across apartment types.

Secondly, the ambiguous impact of apartment supply on net rental yield and apartment prices warrants further research. In my study, I partly rejected my hypothesis that the apartment supply within postal code area would have statistically significant impact on net rental yield. However, due to data limitations, I did have the opportunity to incorporate information regarding the type of constructed apartments, i.e. whether the supply of apartments consisted of one-, two- or three-room apartments. Furthermore, focusing on small geographic area might diminish the impact apartment supply has on macro level. Hence, further studies focusing solely on the impact apartment supply has on rental yield could prove fruitful.

Finally, studies incorporating information regarding upcoming renovations and their relationship to rental yield would be of great interest. As discussed in the results chapter, the upcoming renovations are one of the most important factors rental investors can analyze when acquiring new apartments. As the impact of upcoming repairs has been verified to have significant effect on apartment prices, research focusing on the rental side could potentially yield more information regarding the financial implications renovations have on rental yield.

# 8. REFERENCES


Allen, M.T., Springer, T.M. and Waller, N.G., 1995, *Implicit pricing across residential rental submarkets,* The Journal of Real Estate Finance and Economics, Vol. 11, Iss. 2, pp. 137-151

Alonso, W., 1964, *Location and Land Use,* Harvard University Press

Arimah, B.C., 1997, *The determinants of housing tenure choice in Ibadan, Nigeria*, Urban Studies 34, No. 1, pp. 105-124

Backman, K.E., 2015, *Yksinasuvien helsinkiläisten asumispreferenssit,* Master's Thesis, University of Helsinki, Faculty of Science

Benjamin, J.D. and Sirmans, G.S., 1994, *Apartment rent: Rent control and other determinants*, Journal of Property Research, Vol. 11, Iss. 1, pp. 27-50

Berlemann, M. and Freese, J., 2013, *Monetary policy and real estate prices: a disaggregated analysis for Switzerland,* International Economics and Economic Policy, Vol. 10, Iss. 4, pp. 469-490

Berry, J., McGreal, S., Stevenson, S., Young, J. and Webb, J.R., 2003, *Estimation of Apartment Submarkets in Dublin, Ireland*, Journal of Real Estate Research, Vol. 25, No. 2, pp. 159-170

Bischoff, O. and Maennig, W., 2010, *Rental housing market segmentation in Germany according to ownership*, Journal of Property Research, Vol. 28, Iss. 2, pp. 133-149

Black, R.T., Wolverton, M.L., Warden, J.T. and Pittman, R.H., 1997, *Manufacturing versus distribution: Implicit pricing of Real Property Characteristics by Submarket,* The Journal of Real Estate Finance and Economics, Vol. 15, Iss. 3, pp. 271-285

Bourassa, S.C., Hoesli, M. and Peng, V.S., 2003, *Do housing submarkets really matter?,* Journal of Housing Economics, Vol. 12, Iss. 1, pp. 12-28

Carliner, M. and Marya, E., 2016, *Rental Housing: An International Comparison,* Joint Center for Housing Studies of Harvard University

Cassel, E., and Mendelsohn R., 1985, *The Choice of Functional Forms for Hedonic Price Equations: Comment*, Journal of Urban Economics, Iss. 18, No. 2, pp. 135-142

Chichernea, D., Miller, N., Fisher, J., Sklarz, M., and White, B., 2008, *A Cross-Sectional Analysis of Cap Rates by MSA*, Vol. 30, Iss. 3, pp. 249-292

Chow, G., 1960, *Tests of Equality Between Sets of Coefficients in Two Linear Regressions*, Econometrica, Vol. 28, Iss. 3, pp. 591-605

Clapp, J., Pollakowski, H. O., and Lynford, L., 1992, *Intrametropolitan Location and Office Market Dynamics*, Real Estate Economics, Vol. 20, Iss. 2, pp. 229–257

Cropper, M.L., Deck, L.B., and McConnell, K.E., 1988, *On The Choice of Functional Forms for Hedonic Price Equations*, The Review of Economics and Statistics, Iss. 70, No. 4, pp. 668-675

DiPasquale, D., and Wheaton, W.C., 1996, *Urban Economics and Real Estate Markets*, Englewood Cliffs, New Jersey: Prentice Hall

Dunse, N. and Jones, C., 2011, *The existence of office submarkets in cities,* Journal of Property Research, Vol. 19, Iss. 2, pp. 159-182

Einiö, M., Kaustia, M. and Puttonen, V., 2008, *Price setting and the reluctance to realize losses in apartment markets,* Journal of Economic Psychology, Vol. 29, Iss. 1, pp. 19-34

Fehribach, F.A., Rutherford, E.C. and Eakin, M.E., 1993, *An Analysis of the Determinants of Industrial Property Valuation,* Journal of Real Estate Research, Vol. 8, Iss. 3, pp. 365-376

Fuerst, F., 2008, *Office Rent Determinants: A Hedonic Panel Analysis,* SSRN Electronic Journal, January

Fujita, M., 1989, *Urban Economic Theory. Land Use and City Size,* Cambridge University Press

Garner, T.I. and Verbrugge, R., 2009, *Reconciling user costs and rental equivalence: Evidence from the US consumer expenditure survey,* Journal of Housing Economics, Vol. 18, Iss. 3, pp. 172-192

Glaeser, E.L. and Gyourko, J., 2007, *Arbitrage in Housing Markets,* NBER Working Paper No. 13704

Guntermann, K.L. and Norrbin, S., 1987, *Explaining the Variability of Apartment Rents,* Real Estate Economics, Vol. 15, Iss. 4, pp. 281-404

Hagen, D. and Hansen, J., 2010, *Rental Housing and the Natural Vacancy Rate,* Journal of Real Estate Research, Vol. 32, No. 4, pp. 413-433

Häkkänen, L-E., 2015, *Kuluttajien näkemyksiä vuokra-asumisesta – ”ei tää varmaan loppuelämää kestä”*, Master’s Thesis, University of Helsinki, Faculty of Agriculture and Forestry

Halvorsen, R., and Pollakowski, H.O., 1981, *Choice of Functional Form for Hedonic Price Equations*, Journal of Urban Economics, Iss. 10, No. 1, pp. 37-49

Hanink, D.M., Cromley, R.G. and Ebenstein, A.Y., 2010, *Spatial Variation in the Determinants of House Prices and Apartment Rents in China,* Journal of Real Estate Financial Economics, Vol. 45, Iss. 2, pp. 347-363

Hartzell, D., Hekman, J. and Miles, M., 1986, *Diversification Categories in Investment Real Estate,* Journal of Real Estate Research, Vol. 14, Iss. 2, pp. 230-254

Huovari, J., Laakso, S., Luoto, J. and Pekkala, S., 2002, *Asuntomarkkinoiden alueellinen ennuste*, Pellervo Economic Research Institute, Reports No. 185

Huovari, J., Mäki, F.P. and Volk, R., 2006, *Alueellisten asuntomarkkinoiden kehitys vuoteen 2009,* Pellervo Economic Research Institute, Working Papers, No 86

Janssen, C., Soderberg, B., and Zhou, J., 2001, *Robust Estimation of Hedonic Models of Price and Income for Investment Property*, Journal of Property Investment and Finance, Vol. 19, No. 4, pp. 342-360

Joutsiniemi, J., 2011, *West Metro's effect on housing prices*, Master's Thesis, Aalto University, Department of Finance

Juntto, A., 2007, *Tulot ja kulutus 2007: Suomalaisten asumistoiveet ja mahdollisuudet,* Tilastokeskus

Kain, J.F and Quigley, J.M., 1970, *Explaining Metropolitan Housing Quality,* Journal of the American Statistical Association, Vol. 65, Iss. 330, pp. 532-548

KTI Kiinteistötieto Oy, 2017, The Finnish Property market

Laakso, S., 1997*, Urban Housing Prices and the Demand for Housing Characteristics: A Study on Housing Prices and the Willingness to Pay for Housing Characteristics and Local Public Goods in the Helsinki Metropolitan Area,* ETLA, The Research Institute of the Finnish Economy

Lancaster, K.J., 1966, *A new Approach to Consumer Theory*, The Journal of Political Economy, Vol. 74, No. 2, pp. 132-157

Li, M.M. and Brown, H.J., 1980, *Micro-Neighborhood Externalities and Hedonic Housing Prices,* Land Economics, Vol. 56, No. 2, pp. 125-141

Malpezzi, S., 2003, *Hedonic pricing models: a selective and applied review,* In Housing Economics and Public Policy: Essays in Honour of Duncan Maclennan, edited by T O'Sullivan and K Gibbs. Malder, Massachusetts: Blackwell

Melakari, K., 2014, *Hedonic modelling of residential rents in Helsinki,* Master's Thesis, Aalto University, Department of Economics

Michaels, E.G. and Smith, V.K., 1990, *Market segmentation and valuing amenities with hedonic models: The case of hazardous waste sites,* Journal of Urban Economics, Vol 28, Iss. 2, pp. 223-242

Miller, N.G., 1981, *Residential Property Hedonic Pricing Models: A Review,* In C.F. Sirmans, editor, Research in Real Estate. JAI Press, Inc.

Moilanen, I. and Terho, T., 2010, *Cross-sectional Variation of Rental Yields in the Housing Market – Evidence from Finland*, Master´s Thesis, Aalto-University, Department of Finance

Niemi, K., Peltola, R. and Sandström, T., 2011, *Kansallisuusvarallisuustukimus,* KTI, ROTI

Nikola, N., 2011 *The effect of pipe repairs on housing prices,* Master´s Thesis, Aalto-University, Department of Finance

Ogur, J.D., 1973, *Higher Education and Housing: The Impact of Colleges and Universities on Local Rental Housing Markets,* The American Journal of Economics and Sociology, Vol 32, Iss. 4, pp. 387-394

Oikarinen, E., 2007, *Studies on Housing Price Dynamics*, Turku School of Economics Series A, No. 9

Oinonen, P., 2013, *Analysis of expected sale time and behaviour on the Finnish housing market,* Master's Thesis, Aalto University, Department of Finance

Rosen, S., 1974, *Hedonic prices and implicit markets: Product differentiation in pure competition*, Journal of Political Economy, Vol. 82, Iss. 1, pp. 34–55

Schnare, A.B. and Struyk, R.J., 1976, *Segmentation in urban housing markets,* Journal of Urban Economics, Vol. 3, Iss. 2, pp. 146-166

Sherif, R. and Ashraf, G., 2007, *Pricing inter-neighbourhood variation: A case study on the rental apartment market in Greater Cairo*, Property Management, Vol. 25 Iss. 1, pp. 68-79

Sirmans, G. and J. Benjamin, 1991, *Determinants of Market Rent*, Journal of Real Estate Research, Vol. 6, No. 3, pp. 357-380

Tian, C.Y., 2008, *Arbitrage Conditions, CAP rates, and segmentation in the Housing Market: A Micro Study*, Working Paper

Tiao, G.C., and Goldberger, A., 1962, *Testing Equality of Individual Regression Coefficients*, WEBH Paper 6201, University of Wisconsin, Social Systems Research Institute

Tuominen, J., 2014, *Asumistyytyväisyys Helsingin seudulla,* Master's Thesis, University of Helsinki, Faculty of Political Science

Williams, J.E., 1996, *Real Estate Portfolio Diversification and Performance of the Twenty Largest MSAs,* Journal of Real Estate Portfolio Management, Vol. 2, Iss. 1, pp. 19-30

Wolverton, M.L., Hardin III, W.G. and Cheng, P., 1999, *Disaggregation of Local Apartment Markets by Unit Type,* Journal of Real Estate Finance and Economics, Vol. 19, Iss. 3, pp. 243-257

# 9. APPENDIX

**Table 10 - Linear-linear regression on net rental yield**

This table showcases the linear-linear regressions for the apartment subgroups; one-, two- and three-room apartments and the aggregated sample. The dependent variable is net rental yield in percentages. There are 2,395 one-room apartments, 3,927 two-room apartments and 1,129 three-room apartments in my sample. Column on the left shows the independent variables, followed by columns of one-, two- and three-room apartments. Variables Rent year and Postal code area are controlled for all subgroups. Coefficients are presented without parentheses. T-statistics are in parentheses; *** implies significance at the 0.01 level, ** implies significance at the 0.05 level and * implies significance at the 0.1 level.

| **Variable** | **Aggregate sample** | **One-room apartments** | **Two-room apartments** | **Three-room apartments** |
|---|---|---|---|---|
| *Apartment size* | -0.002 | -0.039 | -0.008 | 0.009 |
| | (-3.02***) | (-13.31***) | (-4.05***) | (1.94*) |
| *New apartment* | -0.439 | -0.479 | -0.484 | -0.308 |
| | (-7.23***) | (-3.77***) | (-6.38***) | (-1.85*) |
| *Built 1962-1983* | 0.498 | 0.352 | 0.587 | 0.719 |
| | (14.85***) | (7.18***) | (12.73***) | (5.89***) |
| *Built 1984-2016* | -0.688 | -0.677 | -0.636 | -0.307 |
| | (-15.14***) | (-8.86***) | (-10.82***) | (-1.97**) |
| *Highest floor* | 0.034 | 0.079 | 0.052 | -0.081 |
| | (1.26) | (1.90*) | (1.48) | (-0.92) |
| *Bottom floor* | 0.089 | 0.046 | 0.091 | 0.130 |
| | (3.04***) | (1.06) | (2.32**) | (1.33) |
| *Sauna* | -0.189 | 0.190 | -0.194 | -0.409 |
| | (-4.84***) | (2.00**) | (-4.03***) | (-3.58***) |
| *Balcony* | -0.242 | -0.149 | -0.293 | -0.207 |
| | (-8.35***) | (-3.54***) | (-7.35***) | (-1.68*) |
| *Owned lot* | -0.268 | -0.164 | -0.319 | -0.427 |
| | (-8.20***) | (-3.17***) | (-7.44***) | (-4.11***) |
| *Elevator* | -0.024 | 0.064 | -0.032 | -0.218 |
| | (-1.00) | (1.77*) | (-0.99) | (-2.58***) |
| *Constuction supply* | 0.000 | -0.000 | -0.000 | 0.000 |
| | (0.13) | (-0.09) | (-0.24) | (0.47) |
| *University* | 0.089 | 0.044 | 0.107 | 0.066 |
| | (5.47***) | (1.60) | (4.99***) | (1.33) |
| *Metro* | 0.078 | 0.059 | 0.097 | 0.074 |
| | (4.40***) | (2.01*) | (4.23***) | (1.24) |
| *Railway* | -0.099 | -0.008 | -0.079 | -0.185 |
| | (-6.38***) | (-0.31) | (-3.84***) | (-3.95***) |
| *Rent year* | Control variable | Control variable | Control variable | Control variable |
| *Postal code area* | Control variable | Control variable | Control variable | Control variable |
| Constant | 7.842 | 6.917 | 9.609 | 6.157 |
| | (13.79***) | (12.37***) | (22.58***) | (8.39***) |
| Adjusted R-Squared | 0.5588 | 0.5519 | 0.6086 | 0.5447 |
| n | 7,451 | 2,395 | 3,927 | 1,129 |

## Table 11 – Double-log regression on dwelling prices

This table showcases the double-log regressions for the apartment subgroups; one-, two- and three-room apartments and the aggregated sample. Dependent variable is the dwelling price with transaction tax. There are 2,395 one-room apartments, 3,927 two-room apartments and 1,129 three-room apartments in my sample. Column on the left shows the variables, followed by columns of one-, two- and three-room apartments. Variables Rent year and Postal code area are controlled for all subgroups. Coefficients are presented without parentheses. T-statistics are in parentheses; *** implies significance at the 0.01 level, ** implies significance at the 0.05 level and * implies significance at the 0.1 level.

| **Variable** | **Aggregate sample** | **One-room apartments** | **Two-room apartments** | **Three-room apartments** |
|---|---|---|---|---|
| *Apartment size* | 0.560 | 0.619 | 0.492 | 0.520 |
| | (91.31***) | (42.32***) | (29.38***) | (10.45***) |
| *New apartment* | 0.144 | 0.104 | 0.152 | 0.120 |
| | (14.28***) | (5.06***) | (12.07***) | (4.91***) |
| *Built 1962-1983* | -0.112 | -0.085 | -0.131 | -0.173 |
| | (-19.75***) | (-10.59***) | (-16.83***) | (-8.97***) |
| *Built 1984-2016* | 0.184 | 0.147 | 0.169 | 0.175 |
| | (24.16***) | (12.08***) | (17.15***) | (7.12***) |
| *Highest floor* | -0.003 | -0.014 | -0.002 | 0.024 |
| | (-0.73) | (-2.07***) | (-0.31) | (1.77*) |
| *Bottom floor* | -0.021 | -0.018 | -0.022 | -0.024 |
| | (-4.21***) | (-2.57***) | (-3.33***) | (-1.64) |
| *Sauna* | 0.055 | -0.027 | 0.059 | 0.057 |
| | (8.48***) | (-1.77*) | (7.30***) | (3.32***) |
| *Balcony* | 0.013 | 0.024 | 0.030 | -0.027 |
| | (2.66***) | (3.52***) | (4.55***) | (-1.46) |
| *Owned lot* | 0.058 | 0.029 | 0.072 | 0.087 |
| | (10.78***) | (3.45***) | (10.11***) | (5.56***) |
| *Elevator* | 0.021 | 0.001 | 0.025 | 0.059 |
| | (5.24***) | (0.16) | (4.75***) | (4.63***) |
| *Constuction supply* | 0.001 | 0.001 | 0.002 | 0.001 |
| | (3.44***) | (0.16) | (3.66***) | (0.51) |
| *University* | -0.047 | -0.036 | -0.056 | -0.034 |
| | (-10.39***) | (-5.21***) | (-9.47***) | (-2.41***) |
| *Metro* | -0.039 | -0.026 | -0.051 | -0.013 |
| | (-9.45***) | (-4.36***) | (-9.10***) | (-0.87) |
| *Railway* | -0.004 | -0.018 | -0.004 | 0.018 |
| | (-1.02) | (-3.17***) | (-0.95) | (1.90*) |
| *Rent year* | Control variable | Control variable | Control variable | Control variable |
| *Postal code area* | Control variable | Control variable | Control variable | Control variable |
| Constant | 9.820 | 9.619 | 10.051 | 9.833 |
| | (122.72***) | (106.09***) | (142.74***) | (44.46***) |
| Adjusted R-Squared | 0.808 | 0.796 | 0.795 | 0.805 |
| n | 7,451 | 2,395 | 3,927 | 1,129 |

## Table 12 – Double-log regression on absolute rent

This table showcases the double-log regressions for the apartment subgroups; one-, two- and three-room apartments and the aggregated sample. Dependent variable is the absolute rent in euros. There are 2,395 one-room apartments, 3,927 two-room apartments and 1,129 three-room apartments in my sample. Column on the left shows the variables, followed by columns of one-, two- and three-room apartments. Variables Rent year and Postal code area are controlled for all subgroups. Coefficients are presented without parentheses. T-statistics are in parentheses; *** implies significance at the 0.01 level, ** implies significance at the 0.05 level and * implies significance at the 0.1 level.

| **Variable** | **Aggregate sample** | **One-room apartments** | **Two-room apartments** | **Three-room apartments** |
|---|---|---|---|---|
| *Apartment size* | 0.387 | 0.257 | 0.253 | 0.527 |
| | (68.60***) | (22.35***) | (16.57***) | (10.60***) |
| *New apartment* | 0.007 | 0.031 | 0.071 | 0.069 |
| | (7.78***) | (1.94*) | (6.17***) | (2.72***) |
| *Built 1962-1983* | -0.038 | -0.019 | -0.042 | -0.107 |
| | (-7.25***) | (-3.03***) | (-5.98***) | (-5.58***) |
| *Built 1984-2016* | 0.099 | 0.079 | 0.083 | 0.126 |
| | (14.12***) | (8.25***) | (9.24***) | (5.19***) |
| *Highest floor* | -0.001 | -0.004 | 0.002 | 0.016 |
| | (-0.28) | (-0.83) | (0.30) | (1.23) |
| *Bottom floor* | -0.006 | -0.009 | -0.011 | -0.005 |
| | (-1.42) | (-1.55) | (-1.87*) | (-0.32) |
| *Sauna* | 0.051 | 0.019 | 0.061 | -0.003 |
| | (8.63***) | (1.59) | (8.40***) | (-0.20) |
| *Balcony* | -0.017 | 0.011 | -0.014 | -0.060 |
| | (-3.78***) | (2.01**) | (-2.36***) | (-3.22***) |
| *Owned lot* | 0.026 | 0.019 | 0.038 | 0.028 |
| | (5.23***) | (1.96**) | (5.93***) | (1.52) |
| *Elevator* | 0.026 | 0.016 | 0.032 | 0.044 |
| | (7.11***) | (3.57***) | (6.55***) | (3.42***) |
| *Constuction supply* | 0.001 | -0.001 | 0.002 | 0.001 |
| | (3.29***) | (-1.28) | (3.33***) | (1.28) |
| *University* | -0.024 | -0.016 | -0.029 | -0.023 |
| | (-5.87***) | (-2.96***) | (-5.56***) | (-1.61) |
| *Metro* | -0.025 | -0.016 | -0.030 | 0.014 |
| | (-6.59***) | (-3.44***) | (-5.88***) | (0.93) |
| *Railway* | -0.018 | -0.025 | -0.015 | -0.008 |
| | (-5.76***) | (-5.58***) | (-3.46***) | (-0.89) |
| *Rent year* | Control variable | Control variable | Control variable | Control variable |
| *Postal code area* | Control variable | Control variable | Control variable | Control variable |
| Constant | 5.304 | 5.399 | 5.469 | 4.361 |
| | (60.84***) | (71.97***) | (82.53***) | (19.60***) |
| Adjusted R-Squared | 0.6295 | 0.5996 | 0.5820 | 0.6191 |
| n | 7,451 | 2,395 | 3,927 | 1,129 |

**Table 13 - Impact of postal code areas on net rental yield**

This table showcases the importance of location to net rental yield. This table shows the coefficients from the controlled postal code areas as discussed in Section 3.4. Only postal code areas with significant coefficients are displayed for the sake of clarity. Other coefficients are as reported previously in Table 5. The postal code area variables are showcased on the leftmost column with the city on the right side in parenthesis. The t-values are in the parenthesis; *** implies significance at the 0.01 level, ** implies significance at the 0.05 level and * implies significance at the 0.1 level.

| Postal code | One-room apartments | Two-room apartments | Three-room apartments |
|---|---|---|---|
| *100 area (Helsinki)* | -0.290<br>(-3.81***) | -0.322<br>(-10.02***) | -0.432<br>(-5.56***) |
| *200 area (Helsinki)* | -0.249<br>(-3.32***) | -0.286<br>(-9.89***) | -0.348<br>(-5.07***) |
| *300 area (Helsinki)* | -0.126<br>(-1.71*) | -0.189<br>(-7.22***) | -0.279<br>(-4.26***) |
| *400 area (Helsinki)* | | | -0.151<br>(-2.34***) |
| *500 area (Helsinki)* | | -0.151<br>(-4.86***) | -0.255<br>(-3.37***) |
| *600 area (Helsinki)* | | -0.176<br>(-6.17***) | -0.249<br>(-3.82***) |
| *700 area (Helsinki)* | | | -0.139<br>(-2.25**) |
| *800 area (Helsinki)* | | | |
| *900 area (Helsinki)* | | | |
| *1200 area (Vantaa)* | | 0.122<br>(3.91***) | 0.164<br>(2.13**) |
| *1300 area (Vantaa)* | | | |
| *1400 area (Vantaa)* | 0.151<br>(1.97**) | 0.091<br>(2.97***) | 0.134<br>(2.03**) |
| *1500 area (Vantaa)* | | | -0.142<br>(-2.03**) |
| *1600 area (Vantaa)* | | -0.049<br>(-1.89*) | |
| *1700 area (Vantaa)* | | | |
| *2100 area (Espoo)* | -0.149<br>(-1.96**) | -0.266<br>(-9.57***) | -0.369<br>(-5.82***) |
| *2200 area (Espoo)* | -0.140<br>(-1.89*) | -0.199<br>(-7.93***) | -0.212<br>(-3.46***) |
| *2300 area (Espoo)* | | -0.109<br>(-4.12***) | |
| *2600 area (Espoo)* | | -0.076<br>(-2.89***) | -0.171<br>(-2.77***) |
| *2700 area (Espoo)* | | | |
| *2900 area (Espoo)* | | | |
| Constant | 2.636<br>(27.88***) | 2.247<br>(31.44***) | 1.731<br>(7.79***) |
| Adjusted R-Squared | 0.5797 | 0.6180 | 0.5780 |
| n | 2,395 | 3,927 | 1,129 |

## Table 14 – Impact of socio-economic variables on net rental yield

This table showcases double-log regression results for the aggregated and disaggregated sample, consisting of all one-, two- and three-room apartments, employing double-log regression model. The dependent variable is net rental yield. This sample includes neighborhood variables: unemployment rate and median income on postal code levels. Sample consists of 389 one-room apartments, 725 two-room apartments and 222 three-room apartments. All independent variables are listed in the leftmost column. The rows with a count (# of dummy) are dummy variables. T-statistics are in parentheses; *** implies significance at the 0.01 level, ** implies significance at the 0.05 level and * implies significance at the 0.1 level.

| **Variable** | **Aggregate sample** | **One-room apartments** | **Two-room apartments** | **Three-room apartments** |
|---|---|---|---|---|
| *Apartment size* | -0.080 | -0.248 | -0.125 | 0.043 |
| | (-5.65***) | (-7.57***) | (-3.37***) | (0.41) |
| *New apartment* | -0.034 | 0.088 | -0.077 | 0.004 |
| | (-1.49) | (0.69) | (-2.63***) | (0.08) |
| *Built 1962-1983* | 0.089 | 0.050 | 0.115 | 0.105 |
| | (6.75***) | (2.34**) | (6.32***) | (2.29**) |
| *Built 1984-2016* | -0.087 | -0.101 | -0.074 | -0.034 |
| | (-5.05***) | (-3.30***) | (-3.15***) | (-0.58) |
| *Highest floor* | 0.030 | 0.035 | 0.041 | 0.039 |
| | (2.81***) | (2.00**) | (2.76***) | (1.23) |
| *Bottom floor* | 0.025 | 0.026 | 0.006 | 0.049 |
| | (2.18**) | (1.49) | (0.38) | (1.29) |
| *Sauna* | -0.044 | 0.093 | -0.045 | -0.134 |
| | (-2.98***) | (2.42***) | (-2.40***) | (-3.21***) |
| *Balcony* | -0.019 | -0.017 | -0.038 | 0.021 |
| | (-1.67*) | (-0.93) | (-2.36***) | (0.48) |
| *Owned lot* | -0.056 | 0.005 | -0.061 | -0.108 |
| | (-4.52***) | (0.20) | (-3.74***) | (-2.72***) |
| *Elevator* | -0.008 | 0.011 | -0.017 | 0.003 |
| | (-0.81) | (0.81) | (-1.30) | (0.08) |
| *Unemployment rate* | 0.037 | 0.054 | 0.040 | 0.065 |
| | (1.16) | (0.53) | (0.65) | (1.05) |
| *Median income* | -0.070 | -0.092 | 0.024 | -0.227 |
| | (-0.88) | (-0.32) | (0.15) | (-1.81*) |
| *Constuction supply* | -0.001 | 0.001 | -0.002 | 0.005 |
| | (-0.29) | (0.72) | (-1.57) | (1.61) |
| *University* | 0.041 | 0.038 | 0.037 | 0.043 |
| | (4.01***) | (2.12**) | (2.69***) | (1.22) |
| *Metro* | 0.009 | 0.014 | 0.012 | 0.018 |
| | (0.98) | (1.01) | (0.93) | (0.53) |
| *Railway* | -0.015 | 0.008 | -0.017 | -0.001 |
| | (-1.64) | (0.45) | (-1.41) | (-0.01) |
| *Rent year* | Control variable | Control variable | Control variable | Control variable |
| *Postal code area* | Control variable | Control variable | Control variable | Control variable |
| Constant | 2.915 | 3.424 | 2.306 | 4.176 |
| | (3.62***) | (1.10) | (1.33) | (2.99***) |
| Adjusted R-Squared | 0.5915 | 0.5925 | 0.6279 | 0.6170 |
| n | 1,336 | 389 | 725 | 222 |

**Table 15 - Correlation matrix of key independent variables**

This table showcases the correlation table between the main variables. The asterix (*) implies that the variable is significant at either 0.05 or 0.01 level.

| **Variables** | *Apartment size* | *New apartment* | *Built 1962-1983* | *Built 1984-2016* | *Highest floor* | *Bottom floor* | *Sauna* | *Balcony* | *Owned lot* | *Elevator* | *Constuction supply* | *University* | *Metro* | *Railway* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Apartment size* | 1 | | | | | | | | | | | | | |
| *New apartment* | 0.0683 | 1 | | | | | | | | | | | | |
| *Built 1962-1983* | 0.1618 | -0.1685* | 1 | | | | | | | | | | | |
| *Built 1984-2016* | 0.1636 | 0.3508* | -0.4805* | 1 | | | | | | | | | | |
| *Highest floor* | 0.0315 | -0.0391* | 0.0770* | -0.0733* | 1 | | | | | | | | | |
| *Bottom floor* | -0.0253 | -0.0157 | 0.0148 | -0.0300* | -0.2471* | 1 | | | | | | | | |
| *Sauna* | 0.2536* | 0.2046* | -0.2864* | 0.6472* | -0.0415* | -0.0138 | 1 | | | | | | | |
| *Balcony* | 0.4549* | 0.1139* | 0.1969* | 0.2943* | 0.0403* | -0.0838* | 0.2328* | 1 | | | | | | |
| *Owned lot* | -0.0599* | 0.0090 | 0.0498* | 0.0255* | -0.0149 | -0.0292* | 0.0498* | -0.0508* | 1 | | | | | |
| *Elevator* | -0.0209 | 0.1701* | -0.1390* | 0.2541* | -0.2003* | -0.1434* | 0.1755* | -0.0008 | 0.1509* | 1 | | | | |
| *Constuction supply* | 0.0634* | 0.1585* | 0.0194 | 0.2291* | -0.0510* | -0.0396* | 0.1580* | 0.1053* | 0.0192 | 0.0896 | 1 | | | |
| *University* | 0.2034* | 0.1057* | 0.3522* | 0.2011* | 0.0612* | 0.0356* | 0.1616* | 0.3465* | 0.0762* | -0.1039* | 0.0414* | 1 | | |
| *Metro* | 0.1906* | 0.1161* | 0.2053* | 0.1953* | 0.0463* | 0.0514* | 0.1501* | 0.2814* | 0.1103* | -0.1047* | 0.1567* | 0.5590* | 1 | |
| *Railway* | 0.0551* | -0.0229* | 0.0807* | -0.0208 | -0.0273* | -0.0022 | -0.0044 | 0.0305* | 0.0058 | 0.0200* | -0.0671* | 0.0109 | -0.3458* | 1 |